前言 PREFACE

为什么要学习博弈论

当我最初决定在大学开设博弈论课程时，首要之务便是定位目标听众。身为一名经济学教授，我希望通过这门课程为经济和工商管理专业的学生提供实用工具，体现其应用价值。此外，最近我对博弈论的研究兴趣已经延伸到经济学范围之外，扩展到竞选、立法、讨价还价等问题上。同时我也在思考：把博弈论应用到政治学中会不会一样有趣？因此，目标听众进一步扩展到了政治学和国际关系学专业。随后，我又想到了博弈论在历史、文学、体育、犯罪、宗教体系、战争、生物学等领域以及在日常生活中的许多有趣的应用，我发现一些初露头角的企业家和政策专家也对这些超越其专业方向的东西感到兴趣盎然。当我仔细审视博弈论在这些不同领域中的各式各样的应用时，我越来越清楚地意识到，博弈论会引发相当多的大学生的兴趣。博弈论是一种适用于所有人类（甚至包括动物王国的某些成员）、会和所有人都不期而遇的推理工具，它应当在普通文科教育中占有一席之地。

经过此番分析，我开始着手构建这门课程（现在已经成文，即本书），它不仅适用于经济学、商学、政治学、国际关系学等博弈理论应用训练的传统学科，还面向大多数普通高校学生。这门课程我已经教了 15 年，但它仍然像我第一次讲授它时那样令我感到新鲜和鼓舞。在将博弈论运用到社会的不同环境、不同领域中的同时，把学生们召集起来集思广益，对形成富有活力、富有洞见的思想论域大有裨益。学生们对这门课程倾注的热情令我惊讶不已。这种热情或许能在课程项目上得到最好的体现，学生分别在真实的、历史的和虚构的世界中运用博弈理论搜寻策略形势并加以分析。课程项目的范围很广，涉及相当多主题，比如伯罗奔尼撒战争（Peloponnesian War）、制药公司间的专利竞争、电视剧《幸存者》（*Survivor*）、会计丑闻、约会困境等。这些项目的完成质量与涉猎程度就是学生们对博弈论兴趣之广、热爱之深的最好见证。这绝对是能点燃学生热情的好课程！

执教大学博弈论课程 15 年后，我明白什么会令人茅塞顿开，什么会令人迷惑不解；我知道什么会让人兴奋不已，什么又会让人昏昏欲睡。正是这些经验（尽管不乏一些让人备感羞辱的经验）为本书提供了大量素材。

PDG

VISIT…

LANZAROTE Caliente.COM

如何学习博弈论

在大学里讲授博弈论课程，在处理学科背景的多样化上会面临一定的挑战。尽管多数学生对学习策略推理有着共同的需求，但他们在数学方面的领悟力却有惊人的差异。我们既要把原料加工好，以确保那些自打高中起就躲避数学的学生也能理解，同时还要保持概念的严谨，以免误导另外一些学得比较好的学生。本书要吸引两类学生，一类是能够不费吹灰之力畅游在数学海洋之中的，还有一类是淹死在海洋中的。第二个挑战则是在传递概念的同时保持学生对该学科的热情。大部分学生都不喜欢深究博弈论的相关概念，但当博弈论被用于解读人类行为时，很少有学生不为之着迷。下面我来说明一下本书是如何应对这些挑战的。

以尽可能少的数学知识为基础

通常每章由一个特定的策略形势开始，吸引读者并引出概念。最先介绍的是用以解决特定情境问题的非正式概念；随后是对概念的系统性分析，依次介绍其构成要件，逐步建立起整体概念或普遍性概念；最终，通过一系列案例固化，增强并进一步激发学生对概念的理解。尽管所运用的数学知识很简单（不会比高中代数更难），但内容并不会因此而打折扣。本书绝不是《博弈论傻瓜手册》或是《策略的傻瓜指南》，而是大量对不完美信息博弈、不完全信息下的信号传递博弈（包括廉价谈话博弈）和远远超越了简单粗暴惩罚的重复博弈的处理。随着结构逐渐搭建起来，本书甚至还提出了经常会令人感到困扰和挫败的更为复杂的条件设置与概念。

从多样化的策略情境展开陈述

多数学生可能来自经济学、商学、政治学和国际关系学等专业，所以来自这些学科的案例在书中占较大比例。尽管如此，这些案例也只不过大约占案例总数的三分之一，因为学生（甚至包括经济学专业的学生）的兴趣远远超越了典型的、传统的博弈论的范围设定。学生们对来自历史、犯罪、体育和日常生活中的案例非常感兴趣（从他们在课程项目中所选择的案例可以看出来），因此本书选取了各种各样的案例，希望能借此保持学生们对本学科的热情。为进一步提升学生的热情，案例通常都伴有丰富的背景介绍，在形式上表现为奇闻轶事（有些是严肃的，有些是好玩的）、引人入胜的旁白、实证数据或者是实验结果。有趣的语境为理论与练习之间建立了关联，为理论的骨架上添加了真实的血肉，让理论鲜活了起来。在本书中，学生不仅能找到解开谜局的聪明答案，更能获得洞察人类行为的能力。

■ **理解测验有助于确保学生对概念的清晰认识。**在每个重要概念后会有一个理解测验，学生可以通过完成理解测验进行自我检测。

■ **将“点拨”用灯泡图标简洁地传达出来。**尽管我们是在特定的策略情境背景下去探索博弈理论，但目的常常是希望能得到一般性的相关结论，我们称之为“点拨”。同时，我们也会将这些分好类的一般性的相关理论运用到博弈论中。

■ **把尚待解决的“谜题”用问号图标凸显出来。**尽管在博弈论的帮助下我们已经能够洞悉大多数人类行为，但还是有许多行为是我们不能理解的。为提醒自己这一事实，也为了将其突出显示给学生，我们把分布在书中各处的、当前尚未解决的、极具挑战性的“谜题”用问号图标凸显出来。

■ **本章总结是对每章要点的综合。**学生们会发现在每章结尾处的本章总结并非仅仅是对关键概念和术语的回顾，还给出了有关本章内容的整体把握。

■ **练习题给学生提供了在各种有趣的情境中运用所学概念和方法解决问题的机会。**有些练习题需要我们重温书中先前提到的案例，有些则引入了新的有意思的情境，其中大部分都源于现实生活中的情况。

PDG

本书的框架结构

现在我将带领你进行一个全书之旅，介绍一下本书的逻辑结构。**第一部分是博弈论概述。**在第 1 章对博弈论进行简要介绍之后，第 2 章是关于如何通过扩展模式和策略模式构建一个博弈的。我的经验是学生对扩展式博弈感到更舒服，原因是其对序列决定的描述可以更容易地映射出真实世界。因此，我从扩展式着手，以一个绑架场景作为开始。策略式紧随其后，与之一道的是一个关于如何在两种模式之间来回往复的讨论。这种表述的优点在于，学生不仅能很快学到可以描述决策结果的策略式博弈，还可以学到扩展式和策略式的关联方式。

尽管与策略情境的模式相比，扩展式显得更加自然，但策略式通常更易于求解。这没什么好奇怪的，因为策略式本来就是作为一种更简洁、更易于控制的数学模型被引入的。我们将在第二部分求解策略式博弈，再在第三部分求解扩展式博弈。

本书第二部分介绍的是求解策略式博弈的方法，其建立在理性行为的基础上，需要树立参与者是理性的这种信念。这不仅仅具备逻辑上的吸引力，更使得学生从简单概念向复杂概念前进了一步。第 3 章从理性参与者假设开始，并将这种假设应用到博弈的求解中。尽管只有特殊的博弈才能单独运用理性参与者假设求解出来，但它给学生提供了一个可以掌握的、用以求解的、最简单的方法。接着我们将这个假定向前推进一点，假设每一位参与者都是理性的，并且每个参与者都相信其他参与者是理性的。这些稍稍加强的假设令我们认识到，并不能单独运用理性参与者假设对博弈求解。我们下一步要做的是，假设每一个参与者都是理性的，每个参与者都相信所有的参与者都是理性的以及每个参与者都相信所有的参与者都相信其他人是理性的。最终我们认为，当理性成为共同知识时，形成了重复剔除严格劣势策略准则（IDSDS）。在第 3 章的附录中，收录了一个更高阶的概念：可理性化策略。尽管有些书把它放在更靠后的位置，但从逻辑上讲此处才是它之所在。原因在于学习了重复剔除严格劣势策略后，学生已具备正确的心态去理解和领会可理性化。

纳什均衡通常是一个对学生来讲更具挑战性的求解概念，因为它要同时解决所有参与者的问题。在第 4 章，我们从一些简单的 2×2 博弈开始，到每人带有三种策略的二人博弈，再到每人带有两种策略的三人博弈。在第 5 章我们会探讨 N 人博弈。在 5.3 节中检验了纳什均衡选择的结果，设计完备。一般读者只需要熟悉纳什均衡（第 4 章所述内容）即可，不必读完第 5 章的剩余部分。第 6 章收录了连续策略博弈，包括“不使用微积分求解纳什均衡”（6.1 节）和进阶课程“运用微积分求解纳什均衡”（6.2 节）。

第二部分的最后讲的是混合策略。第 7 章开始对概率、期望和期望效用理论进行了介绍。考虑到混合策略的复杂性，本章分为几个独立的部分，这样教师可以根据学习的深度从中进行选择。7.1 节至 7.3 节涵盖了基本内容，7.4 节讲的是包含有超过两个以上的参与者或是两种以上策略的更为复杂的博弈，7.5 节讲的是用极大极小策略解决零和博弈。

本书第三部分介绍的是扩展式博弈问题。第 8 章以完美信息博弈开篇，介绍了子博弈精炼纳什均衡的求解概念和逆向归纳法。子博弈精炼纳什均衡的定义接在完美信息博弈之后。这样一来，学生们会觉得这个简单的定义容易理解，好过在第 9 章面对更复杂的适用于不完美信息博弈的定义。在此提供了几个案例，特别要关注等待博弈和消耗战博弈。8.3 节着眼于一些关于逆向归纳法的逻辑性和实验进程方面的争论，提供了可供课堂热烈讨论的话题。第 9 章检验了不完美信息博弈，在介绍了“博弈中的博弈”这一概念以及如何正确地加以分析之后，提出了子博弈精炼纳什均衡的一般性定义。在 9.3 节检验了“承诺”这个概念。

第四部分介绍了不完全信息博弈，这无疑是博弈论课程中最具挑战性的课题。在介绍这一课题时，我采用的方法是在引进这一新概念时放慢速度，用三章的内容讨论这一话题，既有其增量法的实施，也有其覆盖了很多领域的，包括私人信息在内的广泛应用。

第 10 章开始于一个基于 1938 年的《慕尼黑协定》（*Munich Agreement*）的实例，这个实例展示出一个不完美信息博弈如何从一个不完全信息博弈中产生出来，即贝叶斯博弈。随后介绍了贝叶斯纳什均衡的解决办法。第 10 章

专门关注当所有参与者同时行动时，从更细微的信号中摘取出结果。本章从只有一位参与者拥有私人信息的二人博弈开始，然后呈上一个两位参与者都有私人信息的案例。考虑到教师和学生对拍卖都有相当大的兴趣，于是此处案例既包括独立私人价值拍卖、共同价值拍卖，也包括一级密封价格拍卖，并且在本章附录中讨论了类型为连续体时的拍卖博弈。本章附录需要使用微积分，并且是对第 6 章中基于微积分的选读小节的很好补充。（此外，二级密封价格拍卖博弈在第 3 章中介绍了。）

在第 11 章里，假定参与者的行动是连续的，且其中第一个参与者拥有私人信息。然后信号出现，这意味着第二个参与者通过对前者行动的观察，并根据前者的类型更新了自己的信念，做出了第二个贝叶斯均衡。贝叶斯法则及其使用方法在本章的附录中有所介绍。接下来在定义了序贯理性和一致信念的概念之后，提出精炼贝叶斯纳什均衡。这条分析路线一直持续到第 12 章，其关注焦点是廉价谈话博弈。在 12.3 节里，我们会探讨关于表明意图的问题，它和表明信息是对立的。尽管本章没有把不完全信息博弈囊括进来，但表明意图这个问题还是与这一章关于沟通的主题很契合。表明意图部分的素材是对第 9 章的一个有益补充（对第 7 章也是如此），因为在不完美信息博弈中使用了混合策略，且无需借助第四部分用到的其他资料。

第五部分致力于探讨重复博弈问题，因为篇幅足够长，我们可以一步步地接近主题，并对各式各样的应用案例进行深入钻研。第 13 章在第一次世界大战的堑壕战背景下讨论基本机制的传递。在此机制中通过重复博弈可以维持合作。我们为大家展示如何构建重复博弈，并从检验有限次重复博弈入手，但通过检验发现合作无法达成。此时将该博弈扩展为具有无限的或是无穷的范围，才能够确保合作的特征显现出来。本章的关键在于提供了一个可操作的方法去判定某策略组合是否是一个无限次重复的扩展式博弈中的子博弈精炼纳什均衡。该方法基于动态规划并以一种便于使用的形式介绍给大家，其根本含义会在本章的附录中加以介绍。13.4 节针对重复囚徒困境中的合作提供了实验证据，既有实验数据也有来自市场的数据。最后，在附录里描述了如何计算所得收益的现值。

第 14 章与第 15 章通过一系列案例展示了重复博弈的丰富多彩。每个案例都将学生引入一个新的策略情境中，最终得出一个关于如何维系合作机制的普遍结论。第 14 章检验了不同类型的惩罚（例如短暂的严厉惩罚和非对称惩罚）与合作，其中包括轮流帮助对方（互惠型）的合作与非完善监视体系下的合作。第 15 章认为我们的环境很不适合维系合作关系，因为在这种环境中参与者是固定居民，或参与者们罕有互动。尽管如此，在实践中，我们还是能够在如此不适宜的环境中看到合作，第 15 章会为我们展示这种合作是如何实现的。对于固定居民参与者而言，合作可以在重叠交替的世代间得到维系。如果将背景设定为参与者群体会分享信息，那么合作就同样能够在罕有互动的参与者中得到实现。

本书以第六部分的演化博弈理论作为结束。第 16 章的构建围绕着演化稳定策略（ESS）这个概念—— 一种基于找到其余适应值的方法路径（与找出纳什均衡的方法类似），以第 7 章中提及的混合策略为先决条件。第 17 章阐释了一种明确无误的动态研究方法，即使用复制者动态模型（且避免使用混合策略）。第六部分的存在是为了使教师能够兼顾演化稳定策略和复制者动态二者。讲解演化稳定策略时要用到第 16 章的内容。如果专门讲授复制者动态，则需要学生阅读第 16 章的本章导读（导读中对演化博弈理论做了一般性介绍）以及第 17 章（17.3 节除外，该节与解决演化稳定策略的复制者动态下的稳定结果相关）。

关于使用本教材的建议

表 0—1 所示的课程指导是为了帮助教师选择合适的章节。核心一栏包括那些每个博弈论教程都有的章节。广义社会科学一栏基本包括了博弈论应用到社会科学领域时应涉及的所有内容。特别指明一点，由于包括了不完全信息博弈和重复博弈，其内容超出了核心栏。课程指导里还为想要强调私人信息或是反复互动的教师推荐了部分章节。

如果是为某些特定专业开设的课程，如经济学专业或政治学专业，教师可以在核心栏或广义社会科学栏里选择要讨论的概念和相关的应用案例。生物学部分集合了一些本课程的独特元素，研究的是运用博弈论解读动物王国。

设计此课程指导的另一个维度是分析水平。尽管本书写作时设定为面向所有的大学生，但教师仍可以对内容的深浅进行区分。简单栏不涉及概率、微积分（仅在第 6 章和第 10 章附录中出现）和大部分学起来很有挑战性的概念（特别是混合策略与不完全信息博弈）。教师如果对学生有更高要求则可以在进阶栏找到选项，该部分要用到微积分。选择了进阶内容的教师要根据自身的兴趣在各章节中做出选择，比方说，对要用到微积分的较高层次的经济学课程来讲，教师可以将第 10 章（包括附录）、第 11 章、第 13 章以及第 14 章和第 15 章的部分内容增加进来。

表 0—1　课程指导

章节	核心	广义社会科学	私人信息	反复互动	生物学	简单	进阶
1. 策略推理简介	√	√	√	√	√	√	√
2. 构建策略情境模型	√	√	√	√	√	√	√
3. 剔除不可能：理性是共同知识下的博弈求解	√	√	√	√	√	√	√
4. 稳定博弈：两人或三人非连续博弈中的纳什均衡	√	√	√	√	√	√	√
5. 稳定博弈：N人非连续博弈中的纳什均衡		√					√
6. 稳定博弈：连续博弈中的纳什均衡							√
7. 让他们猜不透：随机策略		√	√		√		√
8. 交替：完美信息下的序列博弈	√	√	√	√	√	√	√
9. 黑暗中的交替：不完美信息下的序列博弈	√	√	√	√	√	√	√
10. 我知道你所不知道的信息：拥有私人信息时的博弈		√	√				
11. 你的行为暴露了你的身份：信号传递博弈		√	√				
12. 谎话与说谎者透露的信息：廉价谈话博弈			√				
13. 参与到永远：非固定居所参与者的重复互动博弈		√		√	√	√	
14. 合作和名誉：非固定居所参与者重复互动博弈中的应用		√		√	14.2	√	
15. 非固定居所机构的互动博弈				√			
16. 演化博弈和生物学：演化稳定策略					√		
17. 演化博弈与生物学：复制者动态模型				√	√		

目 录 CONTENTS

第一部分 博弈论概述

第 1 章 策略推理简介 · 3

本章导读

1.1 策略情境采样 · 5
1.2 刺激你的欲望：合并的博弈 · 6
1.3 参与者的心理状况 · 7
1.4 性别代词博弈 · 11

第 2 章 构建策略情境模型 · 12

本章导读

2.1 扩展式博弈：完美信息 · 13
　案例研究 棒球 I · 14
　　伽利略·伽利雷和宗教法庭，I · 15
　　和汽车经销商的讨价还价，I · 16
2.2 扩展式博弈：不完美信息 · 18
　案例研究 抢劫 · 19
　　美国联邦巡回上诉法院 · 20
　　伊拉克战争和大规模杀伤性武器 · 20
2.3 策略是什么 · 23
2.4 策略式博弈 · 25
　案例研究 《托斯卡》 · 25
　　竞选博弈 · 26
　　第 84 期《科学》杂志博弈 · 26
2.5 扩展式转化为策略式 · 26
　案例研究 棒球 II · 27
　　伽利略·伽利雷和宗教法庭，II · 27
　　和汽车经销商的讨价还价，II · 28
2.6 策略式转化为扩展式 · 28

PDG

2.7 共同知识 · 29
2.8 模拟博弈时的其他相关问题 · 31
本章总结与练习题 · 32

第二部分
求解策略式博弈

第 3 章 剔除不可能：理性是共同知识下的博弈求解 · 39

本章导读

3.1 博弈求解：参与者是理性的 · 40
案例研究 白人迁徙与租住房中的种族隔离 · 42
禁止电视播放香烟广告 · 43
eBay 网上的委托出价悖论 · 46
3.2 博弈求解：参与者是理性的和参与者知道参与者是理性的 · 47
案例研究 团队合作 · 48
上帝存在论 · 49
智猪博弈 · 50
3.3 博弈求解：理性成为共同知识 · 51
本章总结与练习题 · 55
附录 1 严格占优策略和弱占优策略 · 58
附录 2 可理性化策略 · 59

第 4 章 稳定博弈：两人或三人非连续博弈中的纳什均衡 · 62

本章导读

4.1 纳什均衡的定义 · 63
4.2 经典的两人博弈 · 65
案例研究 囚徒困境 · 66
协调博弈——行驶规范 · 67
协调和冲突博弈——打电话 · 67
猜谜博弈——石头、剪刀、布 · 68
案例分析 博弈中的冲突和共同利益 · 70
4.3 最优反应法 · 70
4.4 三人博弈 · 71
案例研究 《美国偶像》迷博弈 · 71
投票：真诚的还是虚假的 · 72
晋级与阴谋破坏 · 75
4.5 纳什均衡理论基础 · 77
本章总结与练习题 · 79
附录 纳什均衡的形式化定义 · 81

第 5 章 稳定博弈：N 人非连续博弈中的纳什均衡 · 82

本章导读

5.1 对称博弈 · 83

案例研究 《史尼奇》 · 84

航空安全 · 85

操作系统：苹果还是微软 · 87

实习申请博弈 · 89

5.2 非对称博弈 · 90

案例研究 市场进入博弈 · 91

国内动乱 · 93

5.3 多重纳什均衡的解 · 94

本章总结与练习题 · 97

第 6 章 稳定博弈：连续博弈中的纳什均衡 · 101

本章导读

6.1 不使用微积分求解纳什均衡 · 102

案例研究 相同产品的价格竞争 · 102

价格一致承诺化解价格竞争 · 104

竞选总统 · 106

6.2 运用微积分求解纳什均衡（选读） · 108

案例研究 差别化产品的价格竞争 · 109

公地悲剧——猛犸象的灭绝 · 111

配额资助对慈善捐助的影响力 · 114

本章总结与练习题 · 116

第 7 章 让他们猜不透：随机策略 · 119

本章导读

7.1 在不确定下做决定 · 120

7.2 混合策略和纳什均衡 · 123

7.3 解决混杂合策略纳什均衡问题的提示 · 126

案例研究 第二次世界大战中的阿维让切斯防线 · 127

进入市场 · 129

7.4 进阶提升 · 130

案例研究 足球比赛中的点球 · 130

鞭击：黑色星期五 · 132

旁观者效应 · 133

7.5 纯冲突和防范行为博弈 · 135

本章总结与练习题 · 137

附录 混合策略中纳什均衡的正式定义 · 140

第三部分 求解扩展式博弈

第 8 章 交替：完美信息下的序列博弈 · 143

本章导读

8.1 回顾归纳法及子博弈精炼纳什均衡 · 145
案例研究 古巴导弹危机 · 147
安然公司与检查特权 · 149
种族歧视及运动会 · 150
8.2 等待博弈：抢先博弈与消耗战 · 153
8.3 人类在理智地使用逆向归纳法吗 · 155
本章总结与练习题 · 158

第 9 章 黑暗中的交替：不完美信息下的序列博弈 · 165

本章导读

9.1 子博弈精炼纳什均衡 · 167
案例研究 英国情报机关 · 169
9.2 利用相关均衡收益代替子博弈 · 170
案例研究 OS/2 · 171
参议院的议程决定 · 173
9.3 承诺 · 174
本章总结与练习题 · 179

第四部分 不完全信息博弈

第 10 章 我知道你所不知道的信息：拥有私人信息时的博弈 · 187

本章导读

10.1 不完全信息博弈：《慕尼黑协定》 · 187
10.2 贝叶斯博弈和贝叶斯纳什均衡 · 190
案例研究 狂野西部枪战 · 191
10.3 当所有的参与者都拥有私人信息时：拍卖博弈 · 192
案例研究 独立私人价值和隐藏你的出价 · 193
共同价值和赢者的诅咒 · 194
10.4 委员会和陪审团的投票 · 196
本章总结与练习题 · 200
附录 1 贝叶斯纳什均衡的正式定义 · 203
附录 2 类型为连续体时的一级密封价格拍卖博弈 · 204

第 11 章 你的行为暴露了你的身份：信号传递博弈 · 208

本章导读

11.1 精炼贝叶斯纳什均衡 · 209

案例研究 管理培训生 · 211

11.2 求解信号传递博弈 · 213

案例研究 次品与二手车市场博弈 · 213

恋爱 · 216

边缘政策博弈 · 219

本章总结与练习题 · 221

附录 贝叶斯法则和信念更新 · 224

第 12 章 谎话与说谎者透露的信息：廉价谈话博弈 · 228

本章导读

12.1 博弈论中的交流 · 229

12.2 表明信息 · 230

案例研究 防御性医疗 · 230

购买股票的建议 · 233

12.3 表明意图 · 236

本章总结与练习题 · 240

第五部分

重复博弈

第 13 章 参与到永远：非固定居所参与者的重复互动博弈 · 247

本章导读

13.1 创建一个重复博弈 · 249

13.2 堑壕战：有限次视野 · 251

13.3 堑壕战：无限次视野 · 253

13.4 针对重复“囚徒困境”的实验证据 · 257

本章总结与练习题 · 259

附录 1 所得收益的现值 · 263

附录 2 动态规划 · 266

第 14 章 合作和名誉：非固定居所参与者重复互动博弈中的应用 · 268

本章导读

14.1 惩罚 · 269

14.2 礼尚往来 · 273

案例研究 美国国会和建议经费的花销计划 · 274

吸血蝙蝠及互利主义 · 275

14.3 名誉 · 277

案例研究 做国王的债权人 · 277

福特公司及其日工资 5 美元的规定 · 278
14.4 有缺陷的监视体系和反弹道导弹 · 280
本章总结与练习题 · 282

第 15 章 非固定居所机构的互动博弈 · 286

本章导读

15.1 世代交替间的合作博弈 · 287
案例研究 部落防御 · 287
照顾好你年迈的父母 · 289
执政党和“跛鸭”总统 · 291
15.2 多人合作博弈 · 293
案例研究 eBay · 294
中世纪的商人法 · 296
本章总结与练习题 · 299

第六部分
演化博弈

第 16 章 演化博弈与生物学：演化稳定策略 · 305

本章导读

16.1 鹰鸽博弈 · 306
16.2 演化稳定策略 · 308
案例研究 在牛粪上“生存” · 311
16.3 演化稳定策略的特征 · 312
案例研究 侧边斑点蜥蜴 · 314
16.4 多总体模型 · 316
案例研究 双亲行为 · 316
16.5 敌意的演化 · 318
本章总结与练习题 · 319

第 17 章 演化博弈与生物学：复制者动态模型 · 322

本章导读

17.1 复制者动态和鹰鸽博弈理论 · 323
17.2 复制者动态的一般定义 · 326
17.3 演化稳定策略和复制者动态的吸引子 · 326
案例研究 猎鹿博弈 · 327
棒球中的左右手习惯 · 328
合作的演化 · 331
本章总结与练习题 · 336
理解测验答案 · 338
术语表 · 346
译者后记 · 353

第一部分　博弈论概述

Games, Strategies, and Decision Making

第1章

策略推理简介

学习目标

1.1 策略情境采样

1.2 刺激你的欲望：合并的博弈

1.3 参与者的心理状况

1.4 性别代词博弈

本章导读

> 人们的观念一旦受到一个新的思想影响，就不会再恢复到从前的状态了。
>
> ——奥利弗•温德尔•霍姆斯

2007年4月14日：这些包裹着我的黏糊糊的东西是什么？像这样独自待着，真是无聊。

2007年5月26日：我终于离开这地方了。那个女人为什么对着我笑？我看起来像个邋遢鬼。那个身体扭曲着的、手拿相机的人是谁？

2007年6月1日：哈哈，我得到了。我一哭闹，他们就喂我吃东西。真想知道凭借这招，我还能得到些什么。让我来试一试，如果我吐，会发生什么？哇，我得到了更为体贴细致的照顾。太棒了！

2019年9月24日：来吧，这是一个决定成败的关键性罚球。守门员会向左扑还是向右扑呢？我想我该向右边踢。

2022年1月20日：我必须得到最新款的MP5！它在eBay上的叫价很高，还会再涨价么？如果我只能从eBay上购得最新款的反阻击软件，我是该现在就买呢，还是再等等？

2027年12月15日：这个给我们讲博弈的人太自以为是了。我能准确说出他所有问题的答案。等等，难道这是个诡计？他是不是知道我会那样回答。也许这个人并不像看起来那样木讷。真够狡猾的！

2035年1月7日：如果我想成为销售经理，那么下个季度我的销售业绩就必须名列前茅。但是要做到这一点，我就不能和其他人一样，把目光只局限于老客户身上。也许我得采取点更积极的办法，抓住那些新的大客户。

2071年2月17日：这次调岗简直是逼我离职。也许我可以利用这次机会，好好和老板谈谈我的退休待遇问题。我知道，他很想让我离职。

2089年10月17日：17号房间的那个家伙总是第一个到公共活动室，向人们展示他那愚蠢的脱口秀。他是什么时候成为那里的主角的？明天我宁可少睡一会儿，也要早点赶到那里给他个教训。

从子宫到坟墓，生命就是一系列不断与人交往的过程，包括和我们的父母、亲属、同学、朋友、队友、老师、子女、邻居、同事、领导、咖啡师等。在这本书中，我们会探讨此类相互作用和影响其作用力的两个客观因素。

第一个客观因素是隐藏在人们行为之下的秘密——为什么他们会这样做。假如你是个社会学研究工作者，比如心理学家或经济学家，这也许是你的工作。但对于大多数人而言，这就是他们日常生活的一个组成部分。智人是一种天生有好奇心的物种，特别是当他们彼此靠近时，这一点就表现得尤为突出。关于这个问题，你可以向《人与自然》杂志（*People and National Enquirer*）的编辑们求证。**第二个客观因素就是动机，而引发动机的是需求，而不是好奇心。**与兄弟姐妹的矛盾、参加体育竞赛、参与市场竞争、策划出演一场电视真人秀节目，你可能会遇到许多诸如此类的问题。这时，若能得到一些指导，对你和他人的融洽相处会有很大帮助。

在正文中，我们会探究人们交往中的不同类型，阐明他们之间存在一种**策略性相互依赖**（strategic interdependence）。**什么是策略性相互依赖呢？可以简单概括为：一个人的行为会影响到其他人的利益。**举个例子，一场球赛中，假如是你踢进了制胜一球，不仅你本人会欣喜若狂，你的队友们也会感到欢欣鼓舞，而你的对手会觉得无比沮丧。这个例子说明了人们之间存在相互依赖的关系，而策略性相互依赖则有更丰富的含义。策略性相互依赖是在社会交往中普遍存在的，它是指一个人的最佳决策有赖于他人的行为举动。例如，你会根据对守门员扑球方向的判断，选择是把球踢向球门的左边还是右边。

策略性相互依赖的存在，会对我们的行为选择提出巨大的挑战。假设格雷格和玛西卡到达博物馆后不久就分开了。偏偏这时候格雷格的手机没电了，两人必须独立选择下一步会面的地点。既然格雷格要去他认为玛西卡会去的地方，他就必须像玛西卡那样思考问题。格雷格问自己："假如我是玛西卡，我会去哪里呢？" 但是，当他真的开始这样想时，他马上意识到，玛西卡会到她认为自己会去的地方，也就是说，玛西卡也会思考："假如我是格雷格，我会到哪里去？"因此，格雷格没有必要去想玛西卡会去哪里，他只需要搞清玛西卡会认为自己会去哪里。然而，这件事不会就此停止。就像图 1—1 所描述的那样，一方会思考对方所想的对方所想，这个问题真是太令人厌恶了，于是就有了这样一个名词：**无穷后退**（infinite regress）。

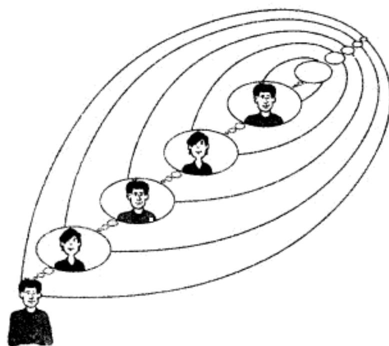


图 1—1　行动的无穷后退

无穷后退是一个只存在于社会科学研究领域中的令人畏惧的概念。它不是产生于物理学、化学，或者其他什么自然科学中。《博弈论与经济行为》（*Theory of Games and Economic Behavior*）是较早的全面阐述该理论的著作之一，作者约翰·冯·诺依曼（John von Neumann）和奥斯卡·摩根斯顿（Oskar Morgenstern）发现了这一有趣的策略形势，同时提出需要有一个全新的工具去应对这一问题：

> 这一社会现象的重要性，显示出的丰富性、多样性以及结构的复杂性，至少可以和物理学相提并论。期待也好，恐惧也好，都需要对这一领域进行更深层面上的精确探究，就如同微积分学一样，然后才能做

出正确的抉择。

博弈论为我们提供了一个可以打破无穷后退锁链的办法，于是我们面对此类问题时，终于可以做些有用的事，不用再头撞南墙了（如果我们的脑袋没有因长时间的撞墙而坏掉）。探究和解读如何发挥博弈论在解决社会问题中的作用是本书的主要任务。

1.1 策略情境采样

从被发现那天起，博弈论就不断在多个领域显示出自身的价值，例如经济、商业、政治和国际关系。本书将会列举出许多成功的案例。同样令人兴奋的是，博弈论还被引入到一些非传统的领域：历史、文学、体育、犯罪、医药、神学、生物和日常生活。为了使读者领略博弈论的应用范围之广，本书从策略角度列举了众多可拓展的案例，下面这些小案例就是本书的一个缩影。

竞价担保。在百思买（Best Buy）进行网上购物时，我们常常会发现这样的声明："假如你发现有同类商品的价格低于百思买商城中的价格，请告诉我们，我们会立刻按此价格结算。"美国最大的电器连锁企业电路城（Circuit City）在其网站上公布了类似的政策，"假如你在本地区其他的商场中发现有同类产品的售价低于我们的价格，请提醒我们。我们很想知道，并愿意为您提供低于该价格10%的折扣价格。"虽然这样言辞凿凿的促销政策表现出极为残酷的市场竞争，但却会造成价格上涨！这是怎么回事？

福特和5美元日薪。1914年，亨利·福特史无前例地给在他汽车工厂工作的工人，每人每天5美元的劳动报酬，这比当时普通工人平均日薪的两倍还多。即便我们可以将之理解为福特先生对待自己的员工宅心仁厚，但这一决策确实给福特汽车公司带来了丰厚的利润回报。为什么高额的人力成本支出，反而使得公司的利润增加呢？

在核问题上的对峙。边缘政策被认为是具有最接近战争而又不真的进入战争的威力。这个技巧，在美国寻求解决朝鲜核计划问题上表现得可谓淋漓尽致。即使金正日无心发动战争，采取一些暗示，表现出要对韩国使用核武器的行动，对他来说也是最为有利的。假如事实如此，布什总统会冒着爆发核战争的风险，摆出强硬的姿态以戳穿金正日的虚张声势吗？

仲裁室。在完成了法庭审问之后，12名陪审员退到仲裁室。在最初的评判中，他们中只有两人认为被告有罪。随后，他们开始审议并依次投票表决。一轮下来，每一个陪审员都投票认定被告有罪。怎么会发生这样的事？如果采用交替投票表决的程序，是否可以避免这样非典型性的事件发生呢？

伽利略和异端审讯。1633年，伟大的天文学家、科学家伽利略被当做异教徒进行审讯。天主教派声称，伽利略的言论违背了教义，不允许他再宣讲日心说。罗马教皇乌尔班一世授意将伽利略的案件提交审判。伽利略应当认罪吗？

登机门等待。一些航班实行坐席开放政策。这意味着，那些排队靠前的乘客可以获得更多选择机会。假如乘客们都觉得坐在门口的座位舒服，那么人们会在什么时候开始排队，你又会选择何时加入？

帮助陌生人。心理学家研究发现，在某些情况下，人们很少愿意对那些需要帮助的人伸出援手，例如，当现场还有其他人在，并且他们完全可以实施帮助时。一些研究甚至发现，可以实施援助的人越多，真正提供帮助的人就越少。为什么人越多，需要帮助的人越容易被忽视呢？

第一次世界大战的堑壕战。第一次世界大战中，交战双方陷入对峙，包括德军在内的同盟军借着防御工事，有规律地发动进攻。在这场血战中，堑壕中敌对的士兵偶然间达成了一

PDG

种休战默契。他们在可预知的时段内开火，以便让另一方接管下一时段进攻的主动。他们不在吃饭时开火，不炮轰对方的补给线。这种休战默契是如何达成并维持的呢？

体育竞赛的兴奋剂。无论是奥运会、全美棒球联盟或是环法自行车赛，服用诸如类固醇之类的违禁药品的后果都是十分严重的。但为什么服用兴奋剂的现象还如此普遍呢？兴奋剂是不可避免的，还是可以被禁止的呢？

猛犸象灭绝。在更新世（Pleistocene era）① 末期，一次大规模的物种消亡席卷了美洲大陆上大多数哺乳动物，这其中就包括猛犸象。这种情形和后来人类出现在地球上的情形很是一致。人类的存在一定会对大自然产生不利的影响吗？这个问题的答案能为解决全球气候变暖问题提供线索吗？

1.2 刺激你的欲望：合并的博弈

博弈理论对于策略情境拓展的意义在于，它开辟出一个更好的了解人们行为的渠道。当问题出现时，博弈理论可以作为我们手中一个有力的武器为我们解开困惑。如果使用得当，我们可以从中学到许多新东西，并且变得更具洞察力。本文将对此理论娓娓道出，从中你会明白这种洞察力是如何获得的。不仅如此，更重要的是，假如你跟着本书思路一路走到最后，你就会培养出自己的这种能力。本节中，我将对博弈论带给我们的洞察力做个简要的展示。

博弈论可以揭示出隐秘的饱含智慧的策略性行为。接下来，我们以大多数人都玩过的一种考验关注力的扑克游戏为例，来让你更清楚地明白其中之意。也许在你曾自觉不自觉地，磕磕绊绊地使用过策略洞察力，下面我们将就此做出详细的描述。博弈论的魅力就在于它能在你采取实际行动之前，给你提供对你所处形势的一种预见性的判断。

这个游戏的规则很简单。所有的52张扑克牌都面朝下排放在桌面上。每个参与者依次选取其中的两张牌。假如这两张牌的点数一致（比如两张都是K），你就可以将其保留并继续选翻。假如它们不一致，你就必须将牌的牌面朝下放回原位，让下一位参与者选取。当所有的牌，即26对牌都被翻取出来后，游戏结束。最终，拿到牌的对数最多的参与者获胜。

赢得这个游戏依靠什么呢？一点点运气，特别是在游戏的开始阶段，我们除了随机翻取之外没有更多的选择。当然，第一个参与者的翻取完全是盲目的，事实上，他拿到一对牌的概率只有6%。但是，随着游戏的深入，记忆力将取代运气成为致胜法宝。当某一张牌被翻起，又因为不能配对而被放回原位时，清楚记下它的位置将会对你下一次配对成功起到帮助。因此，记忆力和运气是这类游戏的宝贵要素（也可统称为运气）。除此之外，当然还有策略。策略？策略在这游戏中扮演什么样的角色呢？

为了聚焦于策略这个维度，我们不考虑记忆力的作用，假设每个参与者都具备很好的记忆。也许你们当中有人像我一样，缺乏这种令人羡慕的特质。出于这一考虑，我们对游戏做了些调整：若两张牌没能被配对，我们将他们面朝上置于桌面。为了不造成混淆，当说到某个参与者选牌时，那张牌可能已经被翻起（以便让所有人都知道那张牌是什么），或者它已经被扣回去（可能它依然面朝上）。

假设有两个参与者安吉拉和扎克正玩这种扑克游戏，扑克牌的摆列如图1—2所示：

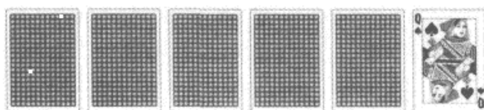


图1—2

图中有6张扑克牌，其中只知道最右边的1张是Q。其余未知的5张牌中，假设有一张Q，两张K，两张10。

现在轮到安吉拉，假设她抽到其中1张为K的扑克牌。具体的情况，如图1—3所示。

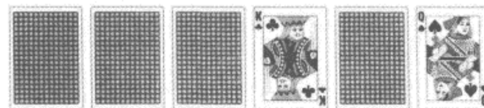


图1—3

①更新世：地质年代名称。——译者注

在此情形下，许多人倾向于翻看另外4张未知的扑克牌，以期望能得到另一张K，而不愿去选择已经明了的那张Q。我们先不急于做这个冒险的举动。假如安吉拉翻取其余的4张未知的牌，那么有1/4的机会可以得到K。因为4张牌中，有1张K，1张Q和2张10。同样的，她有1/4的机会得到Q，有1/2的机会得到10。

如果翻到的是K，会发生什么？那么，安吉拉将得到1对牌，且可以继续翻取。如果翻到的是Q，那么安吉拉没有得到1对牌，选择权将交给扎克，而扎克将面对如图1—4所示的局面。

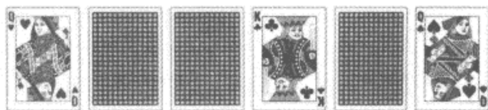


图1—4

因为扎克已经知道两个Q的摆放位置，所以可以很轻松地得到一对牌，如果运气好的话还可以得到更多。最后，假设安吉拉翻开的第二张牌是10，那么轮到扎克时将会是如图1—5所示的局面。

现在扎克将得到全部3对牌！他在面朝下的3张牌中任取1张，都会配成1对牌。假如，他首先翻到的是K，那他只需选第4张牌，就能得到1对K。如法炮制，他就可以得到全部3对牌。

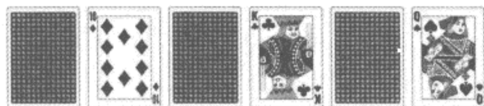


图1—5

表1—1中我们总结了当安吉拉面对图1—3局面时的选择，即她已经翻到了1张K，还需要在4张未知的牌中选出1张作为她的第2张牌。她有25%的机会能配成对（当翻到的牌为K），有25%的机会让扎克拿到1对牌（当翻到的牌的为Q），有50%的机会使得扎克得到所有3对牌（当翻到的牌为10）。

表1—1 当安吉拉得到1张K后，面对的选择及结果

第二张牌	概率	本轮中安吉拉得到牌的对数	下一轮中扎克得到牌的对数
K	25%	1（也许更多）	0（也许更多）
Q	25%	0（确定的）	1（也许更多）
10	50%	0（确定的）	3（确定的）

在任意抽取了1张K之后，安吉拉应该继续随机的抽取下一张牌吗？从博弈论角度推理，安吉拉的最佳策略，不是选择4张未知牌中的任何一张，而是选择已知的Q！解释为什么选Q是最佳策略有偏离本节主题之嫌，但要解释选Q为何最佳却容易。虽然选Q意味着安吉拉将不会配对成功（因为她得到的牌为K和Q），但这也意味着她没有将积极的牌板信息传递给扎克。扎克将面对的局面如图1—6所示。

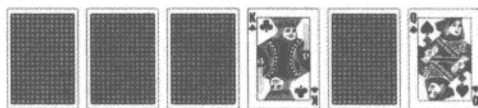


图1—6

我们知道，这种情形下扎克没有确定的把握能配到1对牌。假如换安吉拉选择那未知的4张牌，将有25%的概率，她得到1对牌；有75%的概率，扎克得到至少1对牌。

分析表明了选择一张未知牌的利弊。有利的一面是，它可以使参与者得到一对牌。但为此付出的代价是，若没有选到能配成对的牌（牌依然留在桌面上），宝贵的信息将暴露给你的对手。在特定的情形下，选择一张明知不能配对的牌，以求从策略上限制透露给竞争对手的信息，从而减少其在下一轮中配对的机会，是一种更为明智之选。

一般来讲，博弈论的价值就在于传递诸如此类的洞察力。甚至当我分析一个肯定不现实的模式时——就像我们在翻牌游戏中假定每个参与者都有着完美的记忆力那样，依然可以得到宝贵的经验。在翻牌游戏中，我们要关注的不仅是自己能否配对成功，还要警惕你的行为会给对手传达什么样的信息——特别是那些有利提示，即便他们的记忆力不够完美。

1.3 参与者的心理状况

我认为上帝在创造人类时过高估计了他的能力。

——19世纪英国才子奥斯卡·王尔德

策略情境是由特定环境下人们之间的相互作用决定的。为了便于分析，在进行下一步之前，探讨一下人的定义是很有意义的。假如让你描述一个你熟悉的人，相信会有许多细节浮现在你的脑海，包括这个人的特质、

PDG

智力、学识、头发的颜色、性别、种族、家庭背景、政治信仰、健康状况、个人卫生、音乐品味等。然而，这些关于人的特点，对博弈论而言都可以忽略。在绝大多数情况下，我们要了解或者预测行为只需清楚两点：偏好和信念。

1.3.1 偏好

格瑞丝的手机订制合同到期了，她正在考察两家网络运营商：威瑞森公司（Verizon）和电报电话公司（AT&T）。两家公司提供不同的资费计划和不同型号的手机（其中电报电话公司提供的选择较为诱人，可以支持 iPhone 手机）。**本书中一个关键的假设是，当人们面对选择时，总是可以肯定地说出自己更喜欢哪一个，或者是两者都喜欢。**此假设应用到本节案例中，意味着对于格瑞丝而言，要么威瑞森公司更好，要么电报电话公司更好，要么二者没有分别。诸如格瑞丝这样的人，被称为有完全偏好的人。这样我们就可以将那些因头脑受损造成意志力丧失而不能做出决策的人群排除，对于这部分人，我们将在本书的第二部分谈及。

第二个假设是，一个人的偏好有一定的稳定性和连贯性。举个例子，假如相较威瑞森公司，格瑞丝更喜欢电报电话公司；相较斯普林特公司（Spint），格瑞丝更喜欢威瑞森公司。那么，相较斯普林特公司，她更喜欢电报电话公司。让我们设想一下，如果不是这样，而是相较电报电话公司，格瑞丝更喜欢斯普林特公司，那么格瑞丝的偏好可以总结为：

电报电话公司比威瑞森公司好；
威瑞森公司比斯普林特公司好；
斯普林特公司比电报电话公司好。

让我们看看这样的偏好会给格瑞丝带来什么困惑。

假如格瑞丝从电报电话公司和威瑞森公司开始比较，她将会选择电报电话公司。将电报电话公司和斯普林特公司放在一起，她会想："斯普林特是个不错的选择"。但是，当她准备购买斯普林特的制定服务时，她又想起和威瑞森公司比较下，于是她会觉得威瑞森公司更适合。于是，她又会再次返回，将其和电报电话公司比较，而选择电报电话公司，若再将其和斯普林特公司比较，又会再次选择斯普林特公司。她的比较选择路径构成了一个循环，反复循环下去，她将无所适从。为了避免这样的问题发生，我们假定偏好是具有可传递性的。**偏好可以被传递，假如选 *A* 优于选 *B*，选 *B* 优于选 *C*，那么选 *A* 优于选 *C*。**

偏好如果不可传递，那么造成的后果不仅令人厌烦，反复的举棋不定有时甚至会让人倾家荡产。假设杰克有个不可传递的偏好，选 *A* 优于选 *B*，选 *B* 优于选 *C*，选 *C* 优于选 *A*。假如你有选项 *A*，而杰克有选项 *B* 和选项 *C*。仔细想想表 1—2 中一系列的交换：你提议用你的 *A* 交换杰克的 *B*，外加 1 美元。现在，我们假设杰克喜欢 *A*，并愿意为此付出 *B* 和 1 美元。这样，你就得到了 *B* 和 1 美元，而杰克得到 *C* 和 *A*（同时少了 1 美元）。接下来，你提议用 *B* 换他的 *C* 和 1 美元。因为相较 *B*，杰克更喜欢 *C*（愿意再付出 1 美元），所以杰克将接受交换。

现在你得到 *C* 和 2 美元。下一步，用 *C* 交换 *A* 和 1 美元。既然杰克认为选 *C* 优于选 *A*（假设愿意付出 1 美元），他将进行交易。现在你手上有 *A* 和 3 美元。鉴于我们是以 *A* 和 0 美元起家的，所以和杰克交易真是一本万利的买卖。更棒的是，你可以将这种交易继续下去，且每一轮下来都得到 3 美元。最后，你将会得到杰克的全部金钱。那些没有可传递性偏好的人的人生是多么可悲啊！所以，接受告诫吧，始终保持好自己偏好的可传递性。

表 1—2　榨干杰克的钱

你所有	杰克所有	交换
*A*和0美元	*B*、*C*和99美元	
		*A*交换*B*和1美元
*B*和1美元	*A*、*C*和98美元	
		*B*交换*C*和1美元
*C*和2美元	*A*、*B*和97美元	
		*C*交换*A*和1美元
*A*和3美元	*B*、*C*和96美元	
		*A*交换*B*和1美元
*B*和4美元	*A*、*C*和95美元	
		*B*交换*C*和1美元
*C*和5美元	*A*、*B*和94美元	
⋮	⋮	⋮
*A*和99美元	*B*、*C*和0美元	

假设某人的偏好是完整的且具有可传递性，那么就可将所有的选项赋予特定的数值，以代表这些选项的效用。这样，效用值最高的选项就代表该人的偏好。为了

更加形象地说明这一点，假设有四家公司可供格瑞丝选择，电报电话公司、威瑞森公司、斯普林特公司和 T 移动公司。她的偏好如下：

电报电话公司优于威瑞森公司；
威瑞森公司优于斯普林特公司；
斯普林特公司等于 T 移动公司。

这组偏好意味着如下选择顺序：电报电话公司最优，威瑞森公司次优，斯普林特公司和 T 移动公司同为第三优。下一步，赋予每个选项一个效用值，那么效用最大的选项即为最优方案。具体的效用值如表 1—3 所示。

表 1—3　格瑞斯的效用方程

手机运营服务商	效用
电报电话公司	10
威瑞森公司	6
斯普林特公司	2
T移动公司	2

从表 1—3 中，我们可以清楚地看到能对格瑞斯产生最大效用的选择。就她本人而言，假如所有的四个方案都可行，那么她无疑会选择电报电话公司。因为根据效用最高即为最优方案的原则，电报电话公司的效用值为 10，比效用为 6 的威瑞森公司高，比效用值同为 2 的斯普林特公司和 T 移动公司更高，所以人选。现在，我们假设电报电话公司对于格瑞斯而言是不可行的，那么她就只能从威瑞森公司、斯普林特公司和 T 移动公司中选择了。她的偏好显示威瑞森公司优于其他两个公司，所以她将选购该公司的服务。选择威瑞森公司亦是效用最大法则使然，当她选择范围仅限于威瑞森、斯普林特和 T 移动三家公司时，威瑞森的效用值 6 是最高的。

为了确保效用值最高的选项就是最优方案，选项间的效用值必须予以明确。这样，当且仅当选项 *A* 优于选项 *B* 时，*A* 的效用值大于 *B* 的效用值；当且仅当选项 *A* 无异于选项 *B* 时，*A* 的效用值等于 *B* 的效用值。赋值的方法并不唯一，我们可以将电报电话公司、威瑞森公司、斯普林特公司和 T 移动公司的效用值，由原先的 10，6，2，2，更换为 14，12，11，11，或者 4，3，0，0，凡此种种，遵循其间的比值规律即可。

有关效用的含义没什么深奥的。其主要的功能就是用来描述人们面对一些选择时的偏好。依据偏好的完全性和可传递性，我们可以通过给每个选项赋予一定的数值，进而挑出效用最大的选项，来确定最优方案。如表 1—3 那样，**列出全部选项及其效用值清单，称之为效用函数。一个人的效用函数，将涵盖该人的所有偏好信息。**

某人的偏好，取决于他的成长环境和过程。为什么比起歌剧，这人更喜欢摇滚乐？比起狗更喜欢猫？比起单色的更喜欢条纹的？比起左宗棠鸡更喜欢披萨？决定偏好的因素很多，包括基因、文化、气候、个人经历，还有些说不清的原因。我们并不关心偏好的由来，我们只讨论给定的偏好是如何影响人的行为的。

1.3.2　信念

在很多情况下，效用并非仅仅取决于决策者本人，它还受到其他人的影响。例如，许多手机运营商制定出的资费方案都和用户们所选的网络有某种相互关联，这种关联大多为身处同一运营商网络的手机用户分钟通话费用低于跨网之间的费用。这就意味着，一个人在选择最优服务商时，不得不考虑他的电话常打给谁。

明确了这一点后，具体到格瑞斯的手机业务选择方案中，格瑞斯经常会打给自己最好的朋友丽莎。虽然格瑞斯确实很喜欢电报电话公司（因为她可以得到 iPhone），但和丽莎选择同一个服务商是她必须考虑的。基于此，调整后的格瑞斯的偏好函数如表 1—4 所示。请注意，格瑞斯的最优选择始终等同于丽莎的选择。假如丽莎选用的是斯普林特公司，那么斯普林特对于格瑞斯的效用值就应为最大（参照斯普林特的效用为 6，电报电话为 5，威瑞森为 3）。假如丽莎选择威瑞森，那么威瑞森对于格瑞斯的效用最大。鉴于格瑞斯本人最爱电报电话公司，那么她的效用最大化为，丽莎和她同选电报电话公司。

表 1—4　基于丽莎选择下的格瑞斯的效用函数

格瑞斯所选服务商	丽莎所选服务商	格瑞斯的效用
电报电话公司	电报电话公司	10
电报电话公司	威瑞森公司	5
电报电话公司	斯普林特公司	5
威瑞森公司	电报电话公司	3
威瑞森公司	威瑞森公司	8
威瑞森公司	斯普林特公司	3
斯普林特公司	电报电话公司	2
斯普林特公司	威瑞森公司	2
斯普林特公司	斯普林特公司	6

为了做出最佳的选择，格瑞斯首先必须明确丽莎的

选择。由此推导出参与博弈游戏的人必需具备的第二个特质：对其他参与者可能做出选择的信念。如表 1—4 中描述的那样，虽然我们可以假设人都具有某种偏好，但他们却未必具备信念。事实上，博弈论的一种重要功能就是，使我们获得一种推断别人可能会做什么的信念。

这种信念的产生有两种路径，一种是聪明的路径、一种是愚蠢的路径。愚蠢的路径是一种简单的经验行为，即**体验式学习**（experiential learning）。某人通过不断重复的相互作用，对另一个人可能采取的行为做出正确或错误的判断。这种现象，在自然界中较为普遍，尤其常见于小孩子中和动物王国中。

通过智慧的路径形成的信念，被称为**模拟内省**（simulated introspection）。内省是对自身思想和感受的一种测试，而模拟内省则是一个人通过模拟另一个人的内省过程，形成对其未来行为的判断。虽然经验也是一种路径，但在本书中，模拟内省是一种默认的获得信念的方法。模拟内省是微妙而复杂的，下面我们看看它是如何发挥作用的。

要想获得洞穿别人想法的能力，先得学会**自我认知**（self-awareness），即认识到自身存在。仅仅会思考还够，还要学会思考别人所思考的。思考，不仅是像胃消化食物一样是一个过程，更是一种心理状态。当然，想别人所想未必一定能接近别人真正所想，还需要具有一种被心理学家称为完整的**心智理论机制**（theory of mind mechanism）的要素。具备完整的心智理论机制意味着，你得认识到别人是有思考力的，别人是具备完整的心智理论机制的。也就是说，你得认识到，如同你想别人所想一样，别人有可能也在想你所想。完整的心智理论机制是制定策略的精髓，正是它导致了“无穷后退”的出现。

1.3.3 参与者有何区别

至此，我们讨论了人在哪些方面是相似的：具有符合理论规范的偏好，自我认知和完整的心智理论机制。但是，人们之间又有何不同呢？人身上的三个方面的异质性差别，与博弈理论有特别的联系。首先，虽然假设人人都有完整的、可传递的偏好，但这种偏好却因人而异。例如，格瑞斯喜欢电报电话公司，丽莎喜欢威瑞森公司；托尼是红袜队的球迷，反感扬基队，而强尼却恰恰相反。其次，每个人的选择和机遇不同。比如，在拍卖会上，资金雄厚的竞价者有着其他竞价者不同的选择组合。最后，获取的信心不同。仍拿上个例子来说，资金不充裕的竞价者可能会更全面地评估标的的价值。

值得一提的是技巧。无论是在策略抉择上，还是在计划实施上，都会用到技巧。技巧的构成要素有：天赋、机敏、沉着，还有温斯顿·丘吉尔描述的“天生的、邪恶的诡辩之能，使敌人如被重击般的困惑”。人们具有多好的技巧，他们之间差异又有多大？

图 1—7 中虚构出一组人物智商列表。参与者不会被假定为像爱因斯坦或邦德那样的聪明人，也不会被假定为像克里或憨豆那样的笨人。我们不打算讨论面对核武器对峙时，这三类人会采取什么行动。我们假定参与者至少具有一定的识别力和变通力，即他们智力健全。

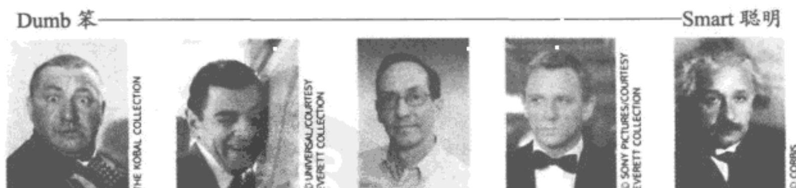


图 1—7　智力排序

人们都可以像星际迷航中的斯波克那样理性么，可以不受情感因素的干扰而做出抉择么？我们的分析是在正常逻辑体系下的推演，不排除人们受情感影响或凭直觉做出决定的可能性。许多情况下，我们不能还原人们是如何做出决定的，然而，我们可以推断某种行为产生的后果。一个人的抉择可能是通过冷静分析得出的，也可能源于自身的经验和直觉。

另一个问题是我们是否允许参与者使用花样繁多的技巧。你能想象海绵宝宝和星球大战中的参议员帕尔帕庭斗智斗勇的场景吗？或者是憨豆先生和伏地魔比拼头脑？尽管这样的猜想会使人兴奋，但我们在此不做讨论。**本书的一个关键假设是，人们都有着相似的技巧。**

策略性的行为限定于一定程度的博弈领域中。虽然，某人可能因拥有更多选择或掌握更多信息而占有优势，但任何人都不拥有高出常人的智慧。（原则上，博弈理论是可以解决此种可能的，但还有待探索。）

1.4 性别代词博弈

在我对博弈理论展开进一步探讨前，要介绍一个关系社会政治的法律性问题：我的出版方、教父以及咖啡店伙计要求我必须向各位读者交代性别代词的用途。任何类似博弈理论等讨论抽象的人的教材，必须明确使用性别代词的意图，并对其用途给出合理的解释。对这类问题的讨论，大多出现在书的前言中，但人们常常忽略对前言的阅读（显然，"前言"在多种语言中就意味着"忽略"）。基于此，有必要在这里让读者们知道，在这种存在争议的问题上我的立场。

本书中涉及的"愚蠢的白人"，就像导演迈克尔·摩尔（Michael Moore）描述的那样，不仅是男人。假如真是这样，那我下一本书中涉及的人将从哪里来？是的，女人也生活在处处充满博弈的世界，她们与男人们同处这个世界。但是在本书的具体案例中如何分配两性比重，却是个悬而未决的问题。我记得一个学者在谈论犯罪时，用男性人称指代罪犯（因为许多罪犯是男性），而用女性人称指代法官、陪审员、律师、证人以及受害者。（补充一点，许多被捕罪犯是男性，也许是因为女性违法后更容易逃脱？）该学者的做法也具有干扰性。那些易受影响的孩子可能会认为他们将会成为罪犯，仅仅因为他们身为男性。我们是否给女孩子们造成了这样的影响——身为女性，面临着被侵害的危险？不同于此类书籍的做法，本书中，那些异常的、排斥社会交往的、性变态的人既可以是男性也可以是女性。

有一个办法就是隐去性别代词，采用能起到替代作用的词汇。在会话交谈中，隐去对方的性别。避免使用"他"、"她"或"他的"、"她的"，以及复数形式"他们"等。在普遍接受异性恋观念的世界，一个同性恋者会尽力避免暴露其伴侣的性别。相似的，在同性恋社区中，一个异性恋者（或是双性恋者）也会把自己隐藏起来。但是，此类中性代词，遇到复数形式时，尴尬就出现了（这会让编辑发疯的）。于是，我想出另一个办法，发明一些中性的代词来取代他、她，它。

经过一段时间的琢磨，参照福克斯电视台的处理办法，我决定以模拟现实世界来解决这个问题。就像人类由男性和女性构成一样，本书中参与博弈游戏的参与者也由男性和女性构成。假如涉及两人博弈，那么就设定一男一女。一般来说，我会将两者穿插使用，这个案例中有男性，那个案例中有女性，实在厌烦了，就举个双性人的例子。必须承认，我没有统计他们出现的频率是否均衡。有兴趣的读者，可以试着统计一下。如果存在不均衡的问题，一旦你将结果告知我，我将很高兴按实际需要，将他们的染色体 X 改成 Y，或将染色体 Y 改成 X。其间，我将尽量保持性别中立。



第2章

构建策略情境模型

学习目标

2.1 扩展式博弈：完美信息
2.2 扩展式博弈：不完美信息
2.3 策略是什么
2.4 策略式博弈
2.5 扩展式转化为策略式
2.6 策略式转化为扩展式
2.7 共同知识
2.8 模拟博弈时的其他相关问题

本章导读

我们是我们无法了解的事物。

——爱默生·皮尤（Emerson Pugh），美国作家

他们说“启示”是这个世界存在的终端产品。迄今为止，对那些尝试理解社会现象的人（如经济学家、政治科学家和吵闹的邻居），或者想要决定如何行为的人（如政策决定者，经营者和青少年）来说，博弈论提供了两点启示。

第一，它提供了一个研究复杂社会现象的框架，且将其简化为可供分析的模型。

第二，它提供了从这些与人们如何采取行动，及人们该如何采取行动相关的模型中总结命题的方法。

本章主要介绍利用博弈论模拟策略情境；下一章是关于如何求解这些博弈的。

人类生存的环境高度复杂，在模拟社会现象时，处理复杂性是一个巨大的挑战。决定将哪些因素加入模型就如同打包行李去大学，你没有办法把所有东西都塞进行李箱。从这方面来说，区分实际模型和虚拟模型就很有必要。实际模型旨在呈现真实世界，是对真实世界的准确描述。和一些情形相比，实际模型可能很复杂。相反，虚拟模型是对真实情境的简化，它并不一定是对真实世界的准确描述。在虚拟模型中，我们尽量从本质上模仿真实世界，而不是复制真实世界。从本质上模仿就是把最关键的因素加入模型，忽略次要因素。本书的大部分模型和为了理解社会现象而构建的模型都是虚拟的。如果分析正确，从虚拟模型中可以得出有关人类如何采取行动的命题。

不论在实际模型还是在虚拟模型中，博弈论都提供了构建模型的“材

PDG

料”。本章主要讲解“材料”的两种类型。扩展式是对策略情境中参与者面临的序列选择的描述。在 2.1 小节，我们思考完美信息博弈。在完美信息博弈中，参与者总是知道博弈中的信息。在 2.2 小节中的模型中，参与者缺乏其他参与者选择策略的信息。2.2 小节还介绍了核心概念——策略，这个概念提供了描述 2.3 小节中提到的策略式的方法，即第二类“材料”。尽管比扩展式更抽象，但策略式更简洁，更便于分析。共同知识是和模拟策略情境相关的概念，也是 2.4 小节的内容。共同知识是所有参与者都知道的信息。

本章的主要内容是构建博弈模型。求解博弈是下一章的内容。

2.1 扩展式博弈：完美信息

博弈论解决的可能是非常严谨的问题，其中一个和绑架相关。绑架在某些国家一直以来都是很严重的社会问题。构建绑架博弈模型可能涉及许多因素。我们的任务不是探索绑架发生的原因，而是学习如何构建博弈理论模型。

因为博弈论的主要目的在于推导人们行为背后的动机，所以一个模型就应该侧重体现个体所做的决定。相应地，我们主要关注绑架者（我们称之为盖伊）和受害人的妻子（我们称之为薇薇卡，她被通知交付赎金）。尽管受害人（我们称之为奥兰多）肯定受信息的影响，但我们假定他不能采取任何选择。在描述这个情境时，我们的模型应该涉及以下几个问题：盖伊和薇薇卡什么时候采取行动？当他们行动时有几种可供选择的行动方案？当他们采取行动时知道的信息是什么？要推导每个参与者未来的行为需要更多的信息，但对初学者来说，回答这些问题得到的信息就足够了。

呈现的模型被称为决策树，如图 2—1。分析决策树是由上而下的（也可以是由左向右的），每个点被称为决策节点，代表博弈中在某个点某些人需要作出选择。从决策节点延伸出来的线是一系列分枝，每个分枝代表可供决策者选择的一种行动方案。选择一个分枝就等于选择了一种行动方案。

在决策树的顶端，盖伊首先做出决定，他的选择是绑架（奥兰多）和不绑架。如果他选择不绑架，那么“博弈结束”。相反，如果他选择绑架奥兰多，那么薇薇卡得知绑架的消息，然后决定是否交付赎金。在等待薇薇卡做出决定时，盖伊决定是释放奥兰多还是杀掉奥兰多。假设在做出选择之前盖伊可以观测到薇薇卡是否交付了赎金（下一章将讨论同时交换的问题）。

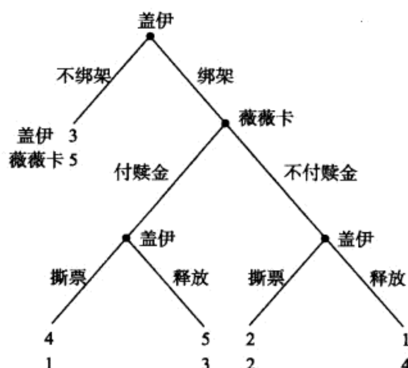


图 2—1　绑架博弈的扩展式

这个博弈有五种结果：每种结果对应决策树的一个路径，或者说对应一系列的行动方案。表 2—1 列出了这五种结果。没有发生绑架事件是一个结果；如果盖伊实施了绑架，有四种可能的结果，分别取决于薇薇卡是否交付赎金以及奥兰多是被释放还是被撕票。

表 2—1　绑架博弈和各个参与者得到的收益

博弈的五种结果	盖伊	（暴力）盖伊	薇薇卡
盖伊不实施绑架	3	3	5
盖伊实施了绑架，薇薇卡交付了赎金，奥兰多被撕票	4	5	1
盖伊实施了绑架，薇薇卡交付了赎金，奥兰多被释放	5	4	3
盖伊实施了绑架，薇薇卡未交付赎金，奥兰多被撕票	2	2	2
盖伊实施了绑架，薇薇卡未交付赎金，奥兰多被释放	1	1	4

构建模型的目的是对盖伊和薇薇卡将采取什么行动做出预测。尽管下一章才涉及求解博弈，但实际上即使我们知道如何求解这个博弈，这里也缺乏足够的信息，因为

在描述参与者如何行动时，只知道他们可选择的行动方案（比如撕票或释放人质）与他们了解到的信息（比如是否交付了赎金）是不够的，我们还需要知道这些参与者最看重的是什么。什么可以使船漂浮起来？什么可以使铃声响起？什么可以激发你的想象力？你明白了吧！

从博弈的五种结果可以分析每个参与者最看重的是什么。假设盖伊仅仅只是想要得到一笔钱，如果不交付赎金，出于报复他会杀掉奥兰多。那么对于盖伊来说，实施绑架，薇薇卡交付赎金，奥兰多被释放是最佳的结果。对他来说，次好结果是实施绑架，收到赎金，杀了奥兰多。次次好的结果是不绑架奥兰多，因为如果薇薇卡不交付赎金，盖伊就不会冒险去绑架奥兰多。在最后两种结果中，假设他绑架了奥兰多，薇薇卡未交付赎金，于是盖伊杀了奥兰多（假设因未得到赎金而报复）。最坏的结果是绑架了奥兰多，薇薇卡未交付赎金，奥兰多被释放。

为了更准确地描述博弈中盖伊的偏好，我们赋予每个结果一个数值，对参与者来说，数值越大表明结果越好。这些数值被称为收益，收益用来测量一个参与者的状况（如事业状况、福利状况或幸福指数）。例如，最高收益表示最好的结果：盖伊实施了绑架，得到赎金，释放了奥兰多。最坏的结果：盖伊实施了绑架，未得到赎金，释放了奥兰多——他得到最低收益 1。

和之前的假设相反，现在假设盖伊认为如果奥兰多被杀死，他被捕的机会就很小，所以他倾向于选择杀掉奥兰多，而不是释放奥兰多。那么盖伊得到的收益为“（暴力）盖伊”一栏中所示的数值。盖伊的最高收益是绑架奥兰多，得到赎金，杀掉奥兰多。

那么薇薇卡呢？如果她关心奥兰多胜过关心钱，那么对她来说，最好的结果是盖伊没有实施绑架行为，我们假设最高收益为 5。对她来说，最糟糕的结果是盖伊实施绑架行为，她交付了赎金，盖伊杀死了奥兰多，她得到最低收益 1。表 2—1 还列出了博弈的其他几种结果中各个参与者得到的收益。

为确保图 2—1 描述的博弈包含所有的相关信息，图中也列出了各个参与者得到的收益值。每个终端节点对应博弈的一个结果，节点下的数值表示对应的结果中盖伊和薇薇卡得到的收益值：第一行是盖伊得到的收益值，第二行是薇薇卡得到的收益值。我们还可以列出奥兰多得到的收益，但这是无关的信息。因为我们的目的是说明人们如何采取行动，而在这个绑架模型中，我们只需要分析绑架者和受害者亲人如何采取行动，只有他们得到多少收益才是至关重要的。

对每种结果只赋予一个收益值这一步类似于第 1 章的赋予选项效用值。在第 1 章，我们研究了个人对特定产品（以手机供应商为例）的偏好，通过效用函数总结了个人的偏好，这里用相同的方法研究个人的偏好。尽管在博弈中应称之为收益，但我们仍可以借用研究效用的方法研究博弈中参与者的偏好。

图 2—1 描述的是扩展式博弈。扩展式博弈包含决策节点，用包含分枝和终端节点的决策树来描述。一个决策节点就是在一个点上，某个参与者需要做出选择。假设图 2—1 包含了所有相关信息：需要做出选择的参与者（盖伊和薇薇卡）；他们采取行动的先后顺序（盖伊首先选择，其次薇薇卡，然后盖伊再次选择）；可供每个参与者选择的行动方案，以及他们如何估算从每种结果中得到的收益值。在这个扩展式博弈中，有 4 个决策节点：在首节点上，盖伊决定是否绑架奥兰多，盖伊的两个决策节点分别表示是撕票还是释放奥兰多（分别对应薇薇卡不交付赎金和薇薇卡交付赎金）。由每个决策节点延伸的有向线被称为分枝。分枝表示在决策点上可供参与者选择的行动方案，更多的分枝意味着更多的选择。

我们称决策树的第一个决策节点为**首节点**（initial node，即博弈开始）；称对应博弈结束的点为**终端节点**（terminal node）。因为该博弈有五种可能的结果，所以在这个博弈中有 5 个终端节点，终端节点与决策节点不同，在终端节点没有参与者采取行动。在终端节点，我们列出了所有参与者得到的收益值。收益值越高表明结果对参与者越有利。

案例研究

棒球 I

棒球比赛中一个众所周知的事实是：一般情况下，右打击球手的表现优于左打投手的表现，左打击球手的表现优于右打投手的表现。表 2—2 解释了这一事实。如果你对棒球比赛的规则不熟悉，那你只需记住安打率等于选手击出的安打数除以打数（换句话说，等于击球手的成功率）。在和右打投手比赛时，右打击球手的成功率是 25.5%，用棒球术语来说即安打率是 0.255；而在和左打投手比赛时，其安打率会高一点，是 0.274。对左打

击球手来说，情况是一样的：在和左打投手比赛时，左打击球手的安打率是 0.266，而在和右打投手比赛时，其安打率出奇的高，是 0.291。现在让我们来探索棒球比赛中一个普遍且简单的事实：惯用手在比赛中的重大意义。

表 2—2　棒球

击球手	投手	安打率
右击	右投	0.255
右击	左投	0.274
左击	右投	0.291
左击	左投	0.266

现在已经是棒球比赛的第九局，巴尔的摩金莺队和纽约扬基队打成平局。纽约扬基队的投手是马里亚诺•里夫萨，他是右打投手。巴尔的摩金莺队的击球手是洛佩斯，他是右打击球手。巴尔的摩金莺队的总教练正在考虑是否把洛佩斯换成左打击球手杰伊•吉本斯。他更愿意让吉本斯对阵里夫萨，以得到更好的结果。然而，作为回应，纽约扬基队的总教练可以将里夫萨替换成左打投手兰迪•约翰逊。相比吉本斯来说，巴尔的摩金莺队的总教练更愿意让洛佩斯对阵约翰逊。当然，两个队的总教练的偏好相对立。

图 2—2 列出了这个形势的扩展式博弈。巴尔的摩金莺队的总教练先决定是否把洛佩斯替换成左打击球手杰伊•吉本斯。如果他选择替换，那么纽约扬基队的总教练决定是否把里夫萨替换成左打投手兰迪•约翰逊。考虑到他们的偏好，当吉本斯对阵里夫萨时，巴尔的摩金莺队的总教练得到最高收益 3；当吉本斯对打约翰逊时，巴尔的摩金莺队的总教练得到最低收益 1。因为假设每个总教练都只在乎输赢，所以，对纽约扬基队总教练最好的结果就是对巴尔的摩金莺队总教练最坏的结果。对纽约扬基队总教练来说，当吉本斯对打约翰逊时是最好的结果；而当吉本斯对打里夫萨时是最坏的结果。

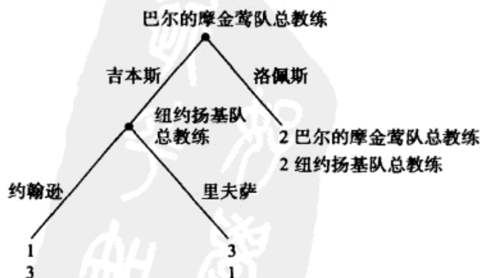


图 2—2　棒球

伽利略•伽利雷和宗教法庭，I

1633 年，伟大的天文学家和科学家伽利略正在考虑是否接受宗教法庭的调查。天主教声称在 1616 年就禁止伽利略宣传和支持哥白尼学说（哥白尼支持日心说），但他却又写出《关于两种世界体系的对话》(*The Dialogue Concerning the Two Chief World System*）来对抗这个禁令。我们模拟这个情境为：天主教堂决定是否把伽利略带到宗教法庭进行审问，如果决定要审问伽利略，那么伽利略决定是否接受审问，而审讯者决定应该问什么问题，采取什么行动。

博弈的参与者是教皇乌尔班八世、伽利略和审讯者（尽管事实上是一个审讯委员会，但这里我们把其合称为一个参与者。）扩展式博弈如图 2—3 所示。乌尔班八世先行决定是否把伽利略的案子移交给宗教法庭。如果他不移交这个案子，那么博弈结束。如果他移交了这个案子，那么伽利略被带到宗教法庭。此时，伽利略必须决定是否承认自己支持哥白尼学说。如果他承认，那么他将受刑且博弈结束；如果他不承认，那么由审讯者决定是否对伽利略施行酷刑。如果审讯者选择不对伽利略施行酷刑，那么从这个意义来看，伽利略赢得博弈，博弈结束。如果审讯者决定对可怜的伽利略施行酷刑，那么伽利略必须决定是否承认。

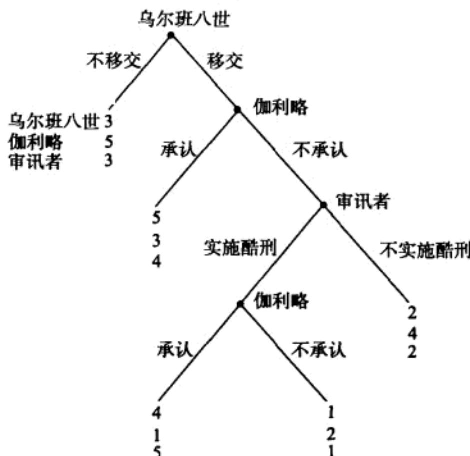


图 2—3　伽利略和宗教法庭

为了完成扩展式博弈，还需要列出各个参与者得到的收益值。这个博弈有 5 个结果：

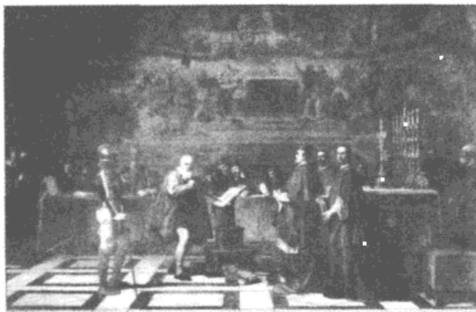
（1）乌尔班八世不移交这个案子；

（2）乌尔班八世移交这个案子，伽利略一开始就承认自己支持哥白尼学说；

（3）乌尔班八世移交这个案子，伽利略一开始不承认自己支持哥白尼学说，被实施酷刑后，承认自己支持哥白尼学说；

（4）乌尔班八世移交这个案子，伽利略一开始不承认自己支持哥白尼学说，被实施酷刑后，依然不承认自己支持哥白尼学说；

（5）乌尔班八世移交这个案子，伽利略一开始不承认自己支持哥白尼学说，审讯者不实施酷刑。



HULTON ARCHIVE/GETTY IMAGES

伽利略和审讯者

在列出收益值时，我们并不是随意给出一个数值，而是给出一个准确反映乌尔班八世、伽利略和审讯者偏好的数值。给出伽利略的收益可能是最简单的工作。他最希望的结果是乌尔班八世不把案子移交给宗教法庭。假设乌尔班八世移交了案子，则伽利略的偏好顺序如下：

（1）他不承认，不被施行酷刑；

（2）他承认；

（3）他一开始不承认，被施行酷刑，仍然不承认；

（4）他一开始不承认，被施行酷刑，最后承认。

伽利略那时已经是一个69岁的老头，证据显示他不会为了坚持原则而接受酷刑。乌尔班八世的情况就有点复杂，因为尽管他希望伽利略承认，但他也不愿意看到一个伟大的科学家被施行酷刑。我们假设乌尔班八世最希望的结果是伽利略承认（不对伽利略施行酷刑），如果结果是伽利略不承认，他宁愿不把案子移交给宗教法庭。审讯者的偏好和乌尔班八世的偏好相似，但他有施虐倾向，更偏好这个结果：对伽利略施行酷刑来逼伽利略承认自己支持哥白尼学说。

那么，伽利略的情况如何呢？等到我们学习了如何求解博弈后，我再给你讲他的收益情况。一旦求解了这个博弈，我再给你讲这段历史知识。

和汽车经销商的讨价还价，I

唐娜想要买辆车，她去了当地一家雷克萨斯汽车店。走进展示厅后，她停留在一辆灰褐色轿车前，售货员马库斯来到她跟前。聊了一会之后，他带她到柜台协商价格。为了简化这个协商过程的模型，假设这辆车有三个可行卖价，分别为p^L，p^M，和p^H，且假设$p^H>p^M>p^L$。（H代表“高”，M代表“中”，L代表“低”）

扩展式博弈如图2—4所示。马库斯先决定给唐娜提出哪一个价钱。唐娜可以接受这个价格（这种情况下，以这个价格成交）或拒绝这个提价。如果她拒绝这个价格，唐娜可以选择离开汽车店（结束谈判），也可以还价。在还价这种情况下，唐娜提高自己的买价，但这样做没有很大作用，所以假设她在比一开始提价低的价格中选择（或拒绝）。比如，如果马库斯提价p^H，那么唐娜可以拿价钱p^L或p^M还价。如果唐娜决定还价，那么马库斯也可以接受或拒绝她的还价。如果他拒绝接受唐娜的还价，那么他可以拿一个稍高的价钱再还价（尽管这个价钱比他开始的提价要低）。这场讨价还价一直持续到唐娜离开，或者协商出一个唐娜和马库斯都可以接受的价钱，或者他们用完这三个价格，还没有达成一个卖价时。

从收益来说，假设如果交易失败，马库斯和唐娜得到的收益都为0。如果交易发生，假设马库斯得到的收益随着汽车卖价的提高而提高，而相应的，唐娜得到的收益降低。更具体地说，如果卖价为p，唐娜得到的收益为p^M-p，马库斯得到的收益为$2(p-p^L)$。（为什么乘以2呢？没有任何特殊的理由。）

如果马库斯以p^L的价钱卖这辆车，那么他得到的收益为0，因为$2(p^L-p^L)=0$。于是他就不在乎是以p^L的价钱卖这辆车还是不卖。若以p^M的价钱卖出，他得到的收益就为正值，这就意味着以p^M的价钱卖出的结果优于不卖出的结果；但如果以p^H的价钱卖出，他得到的收

益更高。对唐娜来说，以 p^M 价钱买入的结果与不买车的结果无异，因为两种情况下她得到的收益都为 0。所以她更愿意以 p^L 的价钱买入，因为这种情况下她得到的收益 $p^M - p^L > 0$，如果以 p^H 的价钱买入，结果更糟，还不如不买，因为这种情况下她得到的收益为 $p^M - p^H < 0$。（是的，收益值也可能是负值。）图 2—4 列出了他们各自得到的收益值。

为了更清晰地阐释这个扩展式博弈，思考当马库斯先行决定提价为 p^H 情况。唐娜可以接受，这样马库斯得到的收益为 2（$p^H - p^L$），唐娜得到的收益为 $p^M - p^H$，或者拒绝。如果拒绝，她可以选择离开或者以价钱 p^M、p^L 还价（记得我们只允许她以低于马库斯的提价还价）。如果唐娜选择以 p^L 还价，那么马库斯可以接受，这样马库斯得到的收益为 0，唐娜得到的收益为 $p^M - p^L$，也可以拒绝。如果拒绝，则马库斯只有一个选择，即以 p^M 还价，唐娜或许接受，或许拒绝。相反，如果唐娜以 p^M 还价，那么马库斯可以接受或拒绝。如果他拒绝，他没有可以还价的机会，博弈结束。

值得注意的是，这个扩展式博弈还可以列出如图 2—5 所描述的博弈。在这个描述中，同一个参与者的决策节点在一排，但并非采取两次行动，而是两个决策节点合并成一个决策节点，这个决策节点包括所有的可行选择方案。比如，在图 2—4 中，唐娜接受或拒绝马库斯的第一次提价 p^M，如果她选择拒绝，她再次做出决定，要么离开，要么以 p^L 还价。同样，当马库斯先行提价为 p^M 时，我们认为唐娜有三个可行方案（分枝）：（1）接受；（2）拒绝且以 p^L 还价；（3）拒绝并离开。图 2—5 和图 2—4 描述的博弈是等同的，因为当求解完博弈时，我们将得到相同的答案。

2.1 理解测验

思考一个包含两个参与者的博弈。在博弈中，父亲可以先行选择是，否，或可能；女儿随后做出选择，她的可选行动方案有：待在家里和去超市。他们各自得到的收益如表 2—3 所示。

表 2—3

结果	父亲得到的收益值	女儿得到的收益值
是，待在家里	8	3
是，去超市	5	9
否，待在家里	4	1
否，去超市	1	5
可能，待在家里	7	2
可能，去超市	2	7

为这个策略式写出相对应的扩展式。

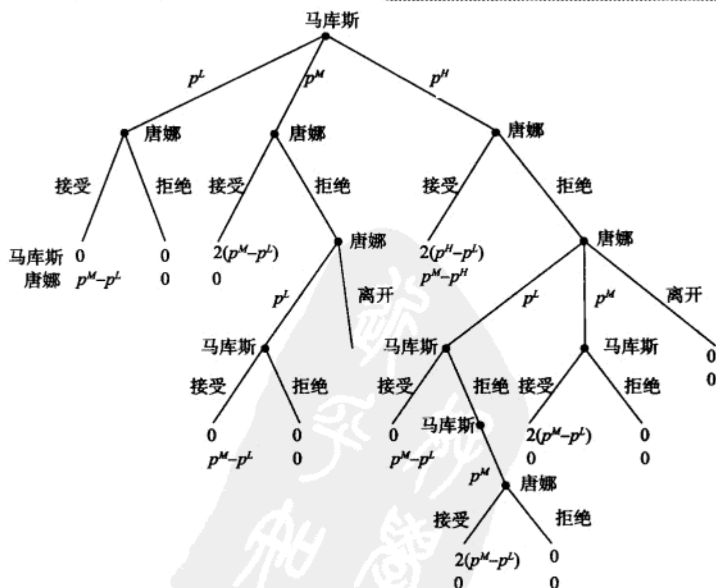


图 2—4　和汽车经销商的讨价还价

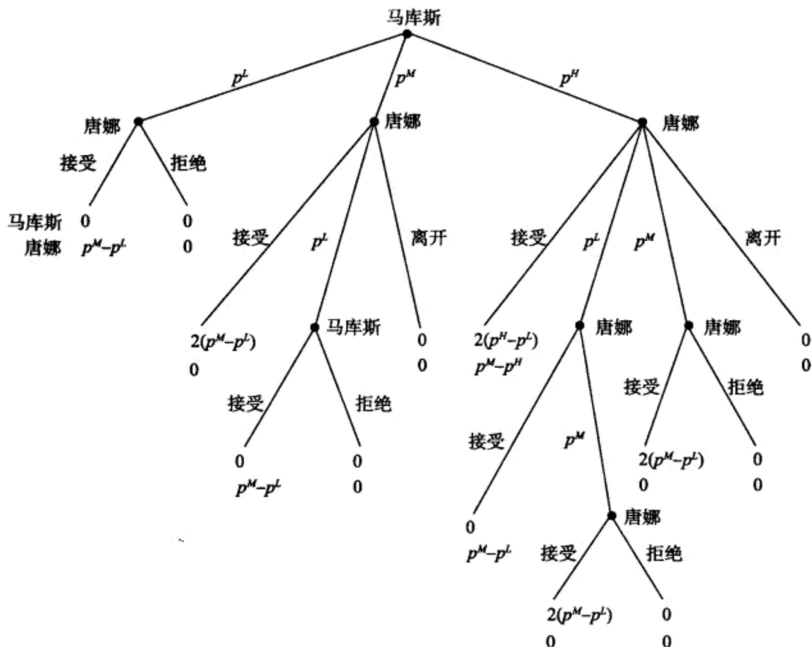


图 2—5 讨价还价博弈扩展式的简化模型

2.2 扩展式博弈：不完美信息

返回绑架博弈。假设在模型中，盖伊（绑架者）和薇薇卡（受害者的亲人）在采取行动时不知道另一个参与者如何行动。图 2—1 描述的扩展式博弈假设盖伊在决定如何处置受害者（奥兰多）时知道薇薇卡是否交付了赎金。另一种情况是在盖伊决定如何处置奥兰多的同时，薇薇卡决定是否交付赎金。你可以想象在城市的某个角落，盖伊在决定是否释放奥兰多，而同时薇薇卡决定是否在规定的地点交付赎金的情形。如何分析这个扩展式博弈？

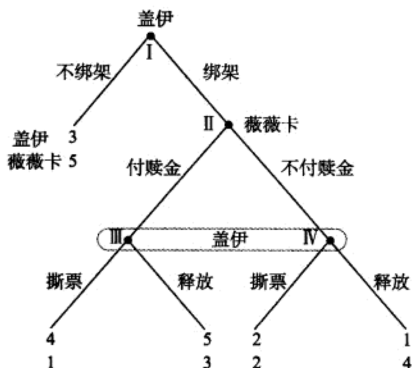
这两个博弈最大的不同在于信息的不同，在图 2—1 中，当盖伊决定是否释放奥兰多时，他知道薇薇卡是否交付了赎金，而当薇薇卡必须决定是否交付赎金时，她不知道盖伊会如何处置奥兰多。现在我们假设当盖伊必须决定是杀掉奥兰多还是释放奥兰多时，他不知道薇薇卡的行动。

下面引入**信息集**（information set）的概念说明这个博弈。一个信息集由所有的决策节点组成。每个决策节点属于且只属于一个信息集。假定一个参与者只知道他处于哪个信息集，不知道其他的如这个信息集有多个节点，那么信息集所属的参与者就不知道该往哪个节点移动。借助例子，我们可以更清楚地阐明这些问题。

图 2—6 是绑架博弈的变形，其中假设当盖伊决定如何处置奥兰多时，不知道薇薇卡是否交付了赎金。从节点和分枝来看，图 2—1 和图 2—6 描述的博弈是相同的，不同之处在于盖伊的两个决策节点——是释放奥兰多还是杀掉奥兰多（用Ⅲ和Ⅳ表示）。当盖伊决定如何处置奥兰多时，这两个决策节点组成了盖伊的信息集。假设盖伊知道在节点Ⅲ和Ⅳ时博弈的情况，但他不知道他所处的是节点Ⅲ还是Ⅳ。想想这意味着什么。不知道他是处在博弈中的节点Ⅲ还是节点Ⅳ意味着盖伊不知道是序列行动“绑架，得到赎金”还是序列行动“绑架，未得到赎金”。这正是我们要构建的模型，在必须决定释放还是杀掉奥兰多时，盖伊不知道薇薇卡是否交付了赎金。所以，盖伊不知道他在博弈中的位置：在节点Ⅲ还是在节点Ⅳ？

在任意一个扩展式博弈中，采取行动的参与者总是拥有一个信息集。这个信息集表示参与者知道的博弈状态。所以当薇薇卡采取行动时，她的信息集是什么呢？她的信息集是节点Ⅱ；换句话说，她准确知道她所处博

弈的阶段。为了与之前一致，我们框住节点Ⅱ表示薇薇卡的信息集。然而，单个节点信息集（只有一个节点的信息集）不设框。在图2—6中，盖伊有两个信息集，一个是首节点（用Ⅰ表示），另一个是包含节点Ⅲ和节点Ⅳ的信息集。



在节点Ⅲ和节点Ⅳ盖伊要决定如何处置奥兰多

图2—6 同时交换的绑架博弈

返回分析薇薇卡，在模型中，她在盖伊决定如何处理奥兰多之前采取行动，当她采取行动时，她不知道奥兰多的情况怎么样。这里，我没有说清采取行动的先后顺序，你注意到了吗？薇薇卡是在盖伊之前还是之后还是同时采取行动？我没有说清楚这个顺序是因为它无关紧要。重要的是信息，而不是在哪天谁做了一个什么决定。关键是薇薇卡决定是否交付赎金时，她不知道盖伊是释放奥兰多还是杀了奥兰多，且当盖伊决定释放还是杀掉奥兰多时，他不知道薇薇卡是否交付了赎金。实际上，图2—7中的扩展式博弈与图2—6中的扩展式博弈相同，只是它颠倒了盖伊和薇薇卡做出决策的顺序。与图2—6中扩展式博弈相同的原因是：我们并没有改变参与者采取行动时的信息集。在这个博弈的两种描述中，我们假设盖伊和薇薇卡同时采取行动，这表明当他们同时采取行动时，拥有的信息集和图2—6中的一样。

图2—1至图2—5中描述的扩展式博弈中，所有的信息集都只有一个节点。这样的博弈被称为**完美信息博弈**（perfect information）。因为当参与者必须做出决定时，他们知道博弈的所有状况。当至少有一个或多个信息集包含多个决策节点时，如图2—6中所描述的博弈，这样的博弈被称为**不完美信息博弈**（imperfect information）。

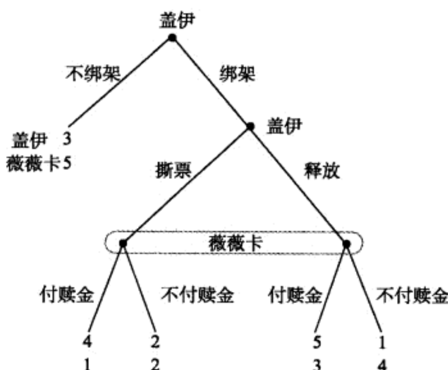


图2—7 同时交换的绑架博弈的扩展式博弈

案例研究

抢劫

> 吝啬的喜剧演员杰克·班尼（Jack Benny）讲述了这样一个故事："我正在一条漆黑的小道上散步，突然从背后出来一个人说，'要钱还是要命？'我站在那儿，全身僵硬。抢劫者又说，'要钱还是要命？'我回答说，'让我想想……让我想想……'"

深夜，西蒙正在回家的路上，突然，他意识到背后有人。在他可以采取任何行动之前，他听到，"我有枪，不要喊，把你的钱包、手机、电脑都给我。"西蒙看不到枪，但他注意到抢匪的手放在外衣口袋里，那里好像有一把枪。如果抢匪口袋中没有枪，西蒙认为他可以狠狠推倒抢匪并成功逃离。但如果抢匪有枪，他的试图逃跑会导致抢匪向他开枪。基于此，他还是宁愿交出他的手机、钱包甚至电脑，而不愿意冒着受伤的危险做逃跑的打算。那晚的早些时候，抢匪一直在想是否带枪，因为如果一个罪犯持枪，被捕后就不仅仅是监狱服刑，所以他宁愿不持枪抢劫。

图2—8中的扩展式博弈描述了以上提及的抢劫情况。抢匪首先从3个选择中做出决定：不携带枪；携带枪，但不向受害者出示枪；携带枪，且向受害者出示枪。针对这三种选择，西蒙必须决定是"猛推抢匪，逃跑"（反抗）或服从抢匪的指令（不反抗）。西蒙有两个信息集，一个是单节点信息集，与抢匪持枪，出示枪相关；

另一个是包含两个节点的信息集，一个是和抢匪携带枪，但不出示枪相关，另一个节点是抢匪没携带枪时的情况。在第二个信息集中，西蒙不确定抢匪口袋中是否装有手枪。

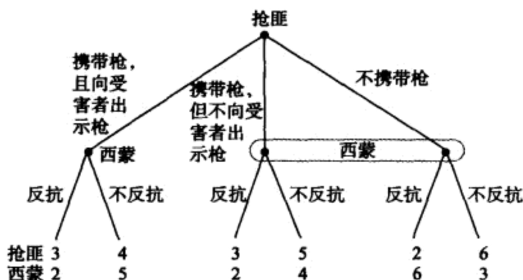


图2—8 抢劫博弈

在具体说明收益情况时，对西蒙来说最好的结果是抢匪没携带手枪，西蒙反抗；最坏的结果是抢匪持枪，西蒙反抗。对抢匪来说，最好的结果是西蒙不反抗，他自己在抢劫中没有携带手枪。最坏的结果是他没有带枪，西蒙反抗，因为在这种情况下，抢匪最终一无所获。

美国联邦巡回上诉法院

当美国联邦巡回法院审理一个案件时，首先从法院的12个陪审员中随机选择3个陪审员组成一个陪审团。在立案后，双方呈交陈述证词的书面材料。如果法庭决定听取口头辩论，那每方律师有15到30分钟的陈述时间。随后3人陪审团审理案件。这里，我们提供一个简单的模型，在模型中，不允许口头辩论。

案件一方的律师是海瑟芙，另一方的律师是约瑟夫。在他们出庭之前，每方律师决定一个法律策略，并根据策略写出一份报告。海瑟芙的策略用策略*A*和策略*B*表示，约瑟夫的策略用策略*I*和策略*II*表示。他们同时提交报告，这就意味着在书写报告时，每个律师都不知道另一方律师写的内容。博弈情形如图2—9所示，其中海瑟芙先采取行动，约瑟夫后采取行动，但在一个信息集中，包含2个节点，海瑟芙选择策略*A*或选择策略*B*。

在读了这两份报告后，法庭的3个陪审员决定投票支持海瑟芙或支持约瑟夫。投票同时进行，每个陪审员在纸上写出自己的投票结果。为了把这个模型简单化，陪审员分别用X，Y和Z表示。如图2—9所示，每个陪审员有4个信息集，每个信息集中都对应律师选择的两个法律策略。陪审员X首先投票，所以他不知道陪审员Y和陪审员Z的投票结果。其次，陪审员Y投票，她不知道陪审员Z的投票结果（因为陪审员Z在她之后投票），但因为信息集的结构她也不知道陪审员X的投票结果。陪审员Y的每个信息集都包含有2个决策节点：一个是陪审员X投票支持海瑟芙的情况，另一个是陪审员X投票支持约瑟夫的情况。再看陪审员Z的信息集，我们发现她的每个信息集包含4个决策节点，分别对应陪审员X和陪审员Y的四种可能的投票结果。尽管描述的是序列投票博弈，实际上每个陪审员投票时都不知道其他两个陪审员的投票结果。换句话说，陪审员同时投票。

伊拉克战争和大规模杀伤性武器

现在我们以较近的一个真实事件为例：当2003年3月20日美国发动对伊拉克战争时，伊拉克、联合国和美国面临的局势。具争议的是萨达姆·侯赛因是否拥有大规模杀伤性武器（WMD）。如图2—10所示，在模型中，伊拉克有两个选择：拥有大规模杀伤性武器，没有大规模杀伤性武器。不知道伊拉克是否有大规模杀伤性武器，联合国决定是否要求检查伊拉克。于是联合国有一个信息集，信息集包含伊拉克的可行方案：一个是拥有大规模杀伤性武器时的方案，另一个是没有大规模杀伤性武器时的方案。如果联合国选择不检查伊拉克，那么美国决定是否入侵伊拉克，这时我们的博弈结束。如果联合国要求检查伊拉克，那么由伊拉克选择行动方案。如果伊拉克没有大规模杀伤性武器，那么它可以选择拒绝检查或接受检查。相反，如果伊拉克拥有大规模杀伤性武器，那么它有3个选择：拒绝检查，接受检查，或接受检查但隐藏大规模杀伤性武器。在最后一个选择中，假设伊拉克成功隐藏了大规模杀伤性武器。假设当伊拉克拥有且不隐藏大规模杀伤性武器时，检查人员发现大规模杀伤性武器。在伊拉克对联合国要求检查的行为采取行动之后，公布检查结果，美国再次决定是否攻打伊拉克。

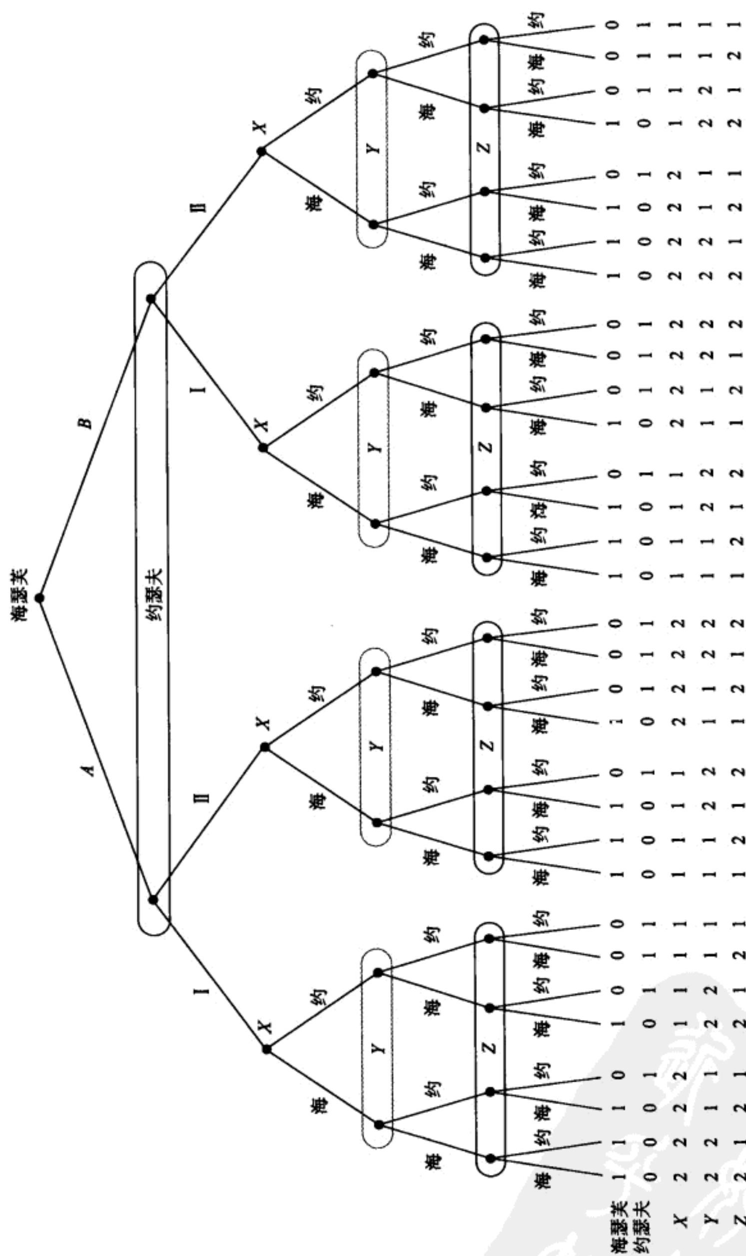


图 2—9　美国联邦巡回上诉法院



图 2—10 伊拉克战争和大规模杀伤性武器

美国有 4 个信息集。用节点 Ⅰ 表示的信息集包含 2 个节点，（1）伊拉克拥有大规模杀伤性武器，联合国选择不检查；（2）伊拉克没有大规模杀伤性武器，联合国选择不检查。尽管美国不能观测到伊拉克的选择，但它可以观测到联合国的决定。信息集 Ⅱ 对应的情况是联合国要求检查伊拉克，伊拉克接受检查，但没发现大规模杀伤性武器（或者是因为伊拉克没有大规模杀伤性武器，或者是因为伊拉克成功隐藏了大规模杀伤性武器从而没有被检查人员发现）。信息集 Ⅲ 表示的情况是联合国要求检查伊拉克，但被伊拉克拒绝，于是，美国不知道伊拉克是否拥有大规模杀伤性武器。对美国来说，唯一的一个单节点信息集是节点 Ⅳ，它是指伊拉克拥有大规模杀伤性武器，联合国要求检查伊拉克，伊拉克接受检查，检查人员发现大规模杀伤性武器的情况。可以做一个类似的练习：描述联合国的一个信息集和伊拉克的 3 个信息集（伊拉克的 3 个信息集都是单节点信息集，因为只有伊拉克隐藏了一些东西）。

2.2 理解测验

现在返回到抢劫博弈中，假设劫匪不仅选择是否携带和出示手枪，还选择是否给枪装上子弹。如果西蒙看到劫匪持枪，他不知道手枪里是否装有子弹。写下这个策略情境下的扩展式博弈（可以忽略得到的收益情况）。

2.3 策略是什么

是故胜兵先胜而后求战，败兵先战而后求胜。善用兵者，修道而保法，故能为胜败之政。

——孙子

你怎样理解孙子的这句名言呢？一个解释是，要成为胜兵，你应该在战争之前详拟计划，然后在战争中执行计划，而不是在战争中企图侥幸取胜。

策略的概念是博弈论的核心，它的定义恰好体现了孙子的思想。**策略是规定参与者在博弈中如何采取行动的决定。**策略不是一系列的行动，而是一系列的应变计划：根据形势决定如何采取行动。如 J. D. 威廉姆斯（J. D. Williams）在一本博弈论书中解释说的，策略是“一个不被对手或自然超越的完美计划。”

我们假设在博弈开始前，参与者选择策略。原则上，这个策略可以是编写的一系列指令，可供另一个参与者参与博弈时选择。换句话说，博弈开始前选择策略就意味着做复杂的脑力活动（运用智力、判断力、才学等）。实际发生的博弈也仅仅是遵循选择策略提供的指令。当然，这个描述有点抽象，因为实际上在策略情境中肯定会用到判断力和智力。然而，如果你要从策略情境中总结命题，你就要接受策略的这一定义。

为了更具体地研究博弈论中策略的本质，我们返回图 2—1 中描述的绑架博弈。绑架者的策略是什么？正如我们刚才所提到的，一个策略是一个决定规定——这个规定描述了每种情形下参与者选择的行动方案，这些行动方案帮助参与者分析自己在博弈中所处的阶段。盖伊（绑架者）发现自己处在 3 种情形下：

（1）考虑是否实施绑架奥兰多（也就是首节点）；
（2）实施绑架奥兰多，拿到薇薇卡交付的赎金，决定释放奥兰多还是杀了奥兰多；
（3）实施绑架奥兰多，未拿到薇薇卡交付的赎金，决定释放奥兰多还是杀了奥兰多。

盖伊发现自己处在 3 种情形和盖伊有 3 个信息集并不是巧合的，一个参与者所处的一种“情形”被定义为一个信息集；所以一种策略赋予一个参与者每个信息集一种行动方案。

那么，盖伊策略的一个模版为：

在首节点，____[填入绑架或不绑架]。
如果实施了绑架，拿到赎金，那么____[填入撕票或释放]。
如果实施了绑架，未拿到赎金，那么____[填入撕票或释放]。

填了这三处空白，可以组合出许多种策略。穷尽所有的可能性，我们得到 8 种可行的策略：

1. 在首节点，绑架。
如果盖伊实施了绑架，得到赎金，那么释放奥兰多。
如果盖伊实施了绑架，但未得到赎金，那么杀了奥兰多。

2. 在首节点，绑架。

如果盖伊实施了绑架，得到赎金，那么释放奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么释放奥兰多。

3. 在首节点，绑架。

如果盖伊实施了绑架，得到赎金，那么杀掉奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么释放奥兰多。

4. 在首节点，绑架。

如果盖伊实施了绑架，得到赎金，那么杀掉奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么杀掉奥兰多。

5. 在首节点，不绑架。

如果盖伊实施了绑架，得到赎金，那么释放奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么杀掉奥兰多。

6. 在首节点，不绑架。

如果盖伊实施了绑架，得到赎金，那么释放奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么释放奥兰多。

7. 在首节点，不绑架。

如果盖伊实施了绑架，得到赎金，那么杀掉奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么释放奥兰多。

8. 在首节点，不绑架。

如果盖伊实施了绑架，得到赎金，那么杀掉奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么杀掉奥兰多。

同理，我们可以定义薇薇卡的策略模版：

如果盖伊实施了绑架，那么选择 ____[填入交付赎金或不交付赎金]。

因为薇薇卡只有一个信息集，她的策略就是一个单一的行动。有一个信息集和两种可行的行动方案，她有两种可行的策略：

1. 如果盖伊实施了绑架，那么选择交付赎金。
2. 如果盖伊实施了绑架，那么选择不交付赎金。

对一个参与者来说，**策略集**（strategy set）是这个参与者所有可行策略的集合。在这个例子中，盖伊的策略集包含 8 种策略，薇薇卡的策略集包含 2 种策略。那么，这个博弈就有 16 种策略组合。

前面提到最困难的决定莫过于选择一种策略，一旦选择了一种策略，参与者开始实施策略，进入博弈。为了更准确地陈述这一点，假设盖伊选择以下策略：

在首节点，实施绑架。

如果盖伊实施了绑架，得到赎金，那么释放奥兰多。

如果盖伊实施了绑架，但未得到赎金，那么杀掉奥兰多。

也假设薇薇卡选择以下策略：

如果盖伊实施了绑架，那么选择交付赎金。

接下来会发生什么呢？根据盖伊的策略，他绑架奥兰多，于是薇薇卡选择交付赎金。得到赎金后，盖伊释放奥兰多。同理，你可以推导 16 种可能策略组合中的任何一个策略组合，以及列出参与者的序列行动。

在你做这个推导之前，注意盖伊一些策略中的特殊情况。8 种策略中的第 5 种到第 8 种规定不绑架，后面又提到如果他选择绑架该如何采取行动。换句话说，策略指导他做一件事，但同时指导他如果他不做他该做的事，该如何采取行动。尽管听起来很奇怪，但我们承认一个参与者策略集包含这些可能的策略有 3 个原因：第一，在定义策略这个概念时，赋予信息集所有可能的策略比剔除这些“可笑”的策略更简单。第二，包含这些可笑的策略不会影响我们得出的结论——命题。第三，也是最重要的一个原因，现在我还不能告诉你。不是我不想告诉你，而是在你理解这个原因之前，你还需要知道一点求解博弈的知识，在第 4 章我会给你一些启示。

> **2.3 理解测验**
>
> 写出图 2—6 描述的绑架博弈中盖伊和薇薇卡的策略集。

2.4 策略式博弈

扩展式博弈是一种常用的博弈问题描述方式。它的优点在于描述了参与者在博弈过程中所遇到决策问题的序列行动，参与者行动时可供其选择的行动方案及所了解的信息集和参与者的支付函数，即博弈结束时每个参与者得到的博弈结果。在这个小节，我们引入另一种描述方式。和扩展式博弈相比，这种描述方式更简单。

策略式博弈（strategic form game，在研究博弈的早期，策略式博弈被称为标准式博弈）包含 3 个要素：

（1）做决策的参与者是谁？
（2）做决策的可选方案是什么？
（3）参与者如何评估不同的决策？

第一个问题的答案是参与者集，第二个问题的答案是参与者的策略集，第三个问题的答案是参与者的收益函数。

参与者集（set of players）指必须做决策的个体集合。决策和策略相关，一个参与者的策略集是他可以选择的策略集合。最后，一个参与者的**收益函数**（payoff function）说明了如何评估一个策略组合（也称策略的集合）。更高的收益意味着参与者得到的博弈结果更好，在求解博弈时，我们将假设每个参与者都努力使自己的收益最大化。

尽管一个参与者在本质上并不重视策略，因为策略仅仅是决策规则，不能吃，不能穿，但一个策略组合决定了博弈的结果（例如：是否会发生绑架案），而参与者在乎博弈结果。在分析之前介绍最后一个专业术语：**多元组**（*n*-tuple），指在一个博弈中策略的组合。多元组包含 *n* 个参与者。

案例研究

《托斯卡》

贾科莫·普契尼（Giacomo Puccini）是最具争议性的伟大歌剧作曲家。在成功创作《波希米亚人》（*La Bohème*）、《蝴蝶夫人》（*Madame Butterfly*）、《图兰朵》（*Turandot*）等著名剧作后，他于 1924 年在布鲁塞尔逝世。

他最成功的剧作之一是《托斯卡》（*Tosca*），这是一部关于爱、忠诚、腐败、淫欲和谋杀的故事——换句话说，这是研究博弈论最完美的材料！剧中主要人物有托斯卡、托斯卡的恋人卡瓦拉多西和警察总监斯卡比亚。斯卡比亚觊觎托斯卡的美貌，设计了一系列毒辣的阴谋。首先，他逮捕了卡瓦拉多西，然后告诉托斯卡第二天早上会将卡瓦拉多西押赴行刑队，他（斯卡比亚）可以令行刑队使用真子弹——这样卡瓦拉多西肯定会死，或者使用空弹——这样卡瓦拉多西可以活下来。斯卡比亚以此要求托斯卡委身于他，托斯卡必须决定是否屈服于他。

在斯卡比亚做出对卡瓦拉多西的判决后，斯卡比亚和托斯卡在晚上见面。托斯卡面对斯卡比亚（她知道斯卡比亚已经做出决定，但不知道决定是什么）要决定是委身于他还是拿刀刺向这个恶魔。

在列出策略式博弈时，我们有两个参与者，斯卡比亚和托斯卡。斯卡比亚的策略集包含两种策略——用真子弹或空弹，而托斯卡可以选择屈服或者刺杀斯卡比亚。如图 2—11 所示，托斯卡的两种策略对应表中两行策略，而斯卡比亚的两种策略对应两列策略。

策略式博弈的最后一个要素是收益。小方格中的第一个数值是托斯卡得到的收益，第二个数值是斯卡比亚得到的收益。比如，如果托斯卡选择刺杀，斯卡比亚选择空弹，那么托斯卡得到收益为 4，斯卡比亚得到的收益为 1。我们已经列出以下四种可能的策略组合（由最好到最坏）：刺杀，空弹；屈服，空弹；刺杀，真子弹；屈服，真子弹。出于对卡瓦拉多西的爱，对托斯卡来说最重要的是斯卡比亚使用空弹。从图中的信息可以看出，对斯卡比亚来说，最优策略组合是（屈服，真子弹），因为在这种情况下，他得到了托斯卡，同时消灭了一个未来竞争托斯卡的对手。对他来说，最劣策略组合是（刺杀，空弹）。

		斯卡比亚	
		真子弹	空弹
托斯卡	刺杀	2, 2	4, 1
	屈服	1, 4	3, 3

图 2—11　托斯卡

图 2—11 被称为**收益矩阵**，它包含了策略式博弈的

所有要素。托斯卡博弈是"囚徒困境"的变形，是博弈论最经典的博弈。在第4章,我们将学习分析"囚徒困境"。

竞选博弈

单词"politics"由"poly"(意思为"许多")和"ticks"(意思为吸血寄生虫)组成。

——拉里•哈迪曼(Larry Hardiman)

思考竞选美国总统的过程。共和党和民主党候选人正在决定他们的竞选纲领。政治光谱为从自由主义到保守主义。我们假设民主党候选人有3种可行的纲领：自由主义、适度自由主义和温和主义。我们也假设共和党有3种可行的纲领：温和主义，适度保守主义和保守主义。

假设一个参与者的收益只取决于候选人的政治意识偏好——他偏好的政治纲领被实施。假设大多数投票者都主张温和主义。假定民主党候选人主张自由主义(也就是说，他最希望实施自由主义)，但他意识到为获得大多数人的选票，他可能要选择一个更温和的竞选纲领。同理，假定共和党候选人主张保守主义，他也知道要选择温和主义来获得更多的选票。收益矩阵如图2—12所示。

		共和党候选人		
		温和主义	适度自由主义	保守主义
民主党候选人	温和主义	4，4	6，3	8，2
	适度自由主义	6，3	5，5	9，4
	自由主义	2，8	4，9	7，7

图2—12　竞选博弈

这些收益反映了两种情况：希望竞选后实施的竞选纲领更接近自己的政治偏好，但同时也希望自己能竞选成功。注意当他/她的对手的政策更激进时，他/她得到的收益就更高，因为对手的政策激进，自己就更容易竞选成功。比如，如果民主党候选人支持适度自由主义纲领，那么当共和党候选人的纲领分别是温和主义、适度保守主义、保守主义时，民主党候选人得到的收益分别为3，5，9。同时注意从策略组合(温和主义，温和主义)到(适度自由主义,适度自由主义)到(自由主义，保守主义)，每个候选人得到的收益都提高了，因为在这三种策略组合中，候选人获胜的概率一样大。

第84期《科学》杂志博弈

第84期《科学》杂志(*Science*)决定为其读者举办一次竞赛：任何参赛者可以要求得到20美元或100美元。如果少于20%的参与者要求得到100美元，那么每个参与者都得到她要求的金额。如果超过20%的参与者要求得到100美元，那么每个参与者都什么也得不到。

参与者集是参加竞赛的所有读者。每个参与者的策略集由3个要素组成：不发送要求，要求得到20美元，要求得到100美元。我们假设每个参与者的收益等于得到的金额减去发送要求的成本，假设发送要求的成本为1美元。

在列出参与者i的收益矩阵时，令x表示选择发送要求得到20美元参与者的人数(不包括参与者i)，y表示选择发送要求得到100美元参与者的人数(不包括参与者i)。那么参与者i的收益函数为：

$$\begin{cases} 0 & \text{如果参与者 } i \text{ 选择不发送要求。} \\ 19 & \text{如果参与者 } i \text{ 发送要求得到 20 美元且} \frac{y}{x+y+1} \leq 0.2 \\ 99 & \text{如果参与者 } i \text{ 发送要求得到 100 美元且} \frac{y+1}{x+y+1} \leq 0.2 \\ -1 & \text{如果参与者 } i \text{ 发送要求得到 20 美元且 } 0.2 < \frac{y}{x+y+1} \\ -1 & \text{如果参与者 } i \text{ 发送要求得到 100 美元且 } 0.2 < \frac{y+1}{x+y+1} \end{cases}$$

例如，如果参与者i发送要求得到20美元，且少于20%的参与者发送要求得到100美元(即$\frac{y}{x+y+1} \leq 0.2$)时，那么她从《科学》杂志得到20美元,减去发送成本1美元，即19美元。

尽管知道博弈的结果有很大意义,但第84期《科学》杂志不会举办这类竞赛，因为保险公司不愿意为出版商的竞赛损失投保。

2.5 扩展式转化为策略式

对每一个扩展式博弈来说，总存在一个策略表达式。这里，我们举以下几个例子来证明如何推导相对应的策略式博弈中的参与者集(最容易推导的一项)，每个参与者的策略集和收益函数。

案例研究

棒球，II

思考图2—2中描述的棒球博弈。巴尔的摩金莺队总教练的策略集包括两个要素：（1）把洛佩斯替换成左打击球手吉本斯；（2）保留洛佩斯。如图2—2所示，对纽约扬基队的总教练来说，只存在一个单节点信息集，所以他的策略也只包含一个行动方案。他的策略集包括（1）将里夫萨替换成左打投手约输逊；（2）保留里夫萨。在构建收益矩阵时，你仅仅需要思考这四个可能的策略组合，然后决定这四个策略组合通向哪个终端节点。

如果策略组合是（保留洛佩斯，保留里夫萨），那么巴尔的摩金莺队的总教练得到的收益为2，纽约扬基队的总教练得到的收益也为2。相反，如果是策略组合（保留洛佩斯，将里夫萨替换成约输逊），博弈的路径和各自得到的收益与采取策略组合（保留洛佩斯，保留里夫萨）时博弈的路径和得到的收益相同，因为将里夫萨替换成约输逊就意味着“如果将洛佩斯替换成吉本斯，就将里夫萨替换成约输逊”。当策略组合是（将洛佩斯替换成吉本斯，保留里夫萨）时，吉本斯的表现优于里夫萨的表现，得到的收益组合为（3，1），其中第一个数值是巴尔的摩金莺队总教练得到的收益值。最后，如果策略组合是（将洛佩斯替换成吉本斯，将里夫萨替换成约输逊），吉本斯的表现优于约输逊的表现，得到的收益组合是（1，3）。于是收益矩阵就如图2—13所示。

纽约扬基队总教练

		保留里夫萨	将里夫萨替换成约输逊
纽约扬基队总教练	保留洛佩斯	2,2	2,2
	将洛佩斯替换成吉本斯	3,1	1,3

图2—13 棒球博弈的策略式

伽利略·伽利雷和宗教法庭，II

返回图2—3中所描述的博弈，我们知道伽利略有两个信息集：一个是与教皇乌尔班八世把案件移交给审讯者相关，另一个是与罗马教皇乌尔班八世把案件移交给宗教法庭后，伽利略不承认，审讯者对伽利略施行酷刑相关。于是，对伽利略来说，一个策略就包括两种行动方案。我们令*C/DNC*（承认/不承认）表示伽利略的策略，即在第一个信息集中选择承认，在第二个信息集中选择不承认。其他三个策略是：*C/C*，*DNC/C*，和*DNC/DNC*。审讯者有一个信息集（当教皇把案件移交且伽利略不承认）和两种行动方案。这使得他有两种策略：施行酷刑和不施行酷刑。乌尔班八世也只有一个信息集（也就是首节点），且因为他可以选择移交案件与不移交案件，所以他有两个策略：移交案件与不移交案件。

如图2—14所示的收益矩阵，横行表示伽利略的策略，竖行表示审讯者的策略。整个矩阵表示乌尔班八世的策略。每一个单元格的第一个数值表示伽利略得到的收益，第二个数值表示审讯者得到的收益，第三个数值表示教皇乌尔班八世得到的收益。思考在填写收益矩阵时，比如在策略组合（移交案件，*DNC/C*，不施行酷刑）中三人各自得到的收益情况。当路径为教皇移交案件，伽利略不承认，审讯者不对伽利略施行酷刑时，得到的收益情况为如图所示的（4，2，2）。注意不论伽利略和审讯者采取何种策略，如果教皇不移交案件，得到的收益情况都为（5，3，3），因为如果教皇不移交案件，伽利略和审讯者就没有机会采取行动。同理，如果教皇移交案件，伽利略第一次承认（或者选择策略*C/C*，或者选择策略*C/DNC*），那么不论审讯者打算施行酷刑还是不打算施行酷刑，得到的收益情况相同，因为伽利略的承认就表明审讯者没有采取行动的机会。

审讯者

		施行酷刑	不施行酷刑
伽利略	C/C	3,4,5	3,4,5
	C/DNC	3,4,5	3,4,5
	DNC/C	1,5,4	4,2,2
	DNC/DNC	2,1,1	4,2,2

教皇乌尔班八世——移交案件

审讯者

		施行酷刑	不施行酷刑
伽利略	C/C	5,3,3	5,3,3
	C/DNC	5,3,3	5,3,3
	DNC/C	5,3,3	5,3,3
	DNC/DNC	5,3,3	5,3,3

教皇乌尔班八世——不移交案件

图2—14 伽利略和审讯者博弈的策略式

和汽车经销商的讨价还价，II

这个模型是迄今为止我们学习到的最复杂的一个博弈模型。思考图 2—5 中描述的博弈。马库斯有四个信息集：(1) 首节点，(2) 与他提价为 p^M，唐娜拒绝接受这个要价且还价为 p^L 相关的节点，(3) 与他提价为 p^H，唐娜拒绝接受这个要价且还价为 p^M 相关的节点，(4) 与他提价为 p^H，唐娜拒绝接受这个要价且还价为 p^L 相关的节点。于是马库斯的策略模版如下：

在首节点，提价为________[填入 p^L，p^M，或 p^H]。

如果我提价为 p^M，唐娜拒绝接受且还价为 p^L，那么______[填入接受或拒绝]。

如果我提价为 p^H，唐娜拒绝接受且还价为 p^M，那么______[填入接受或拒绝]。

如果我提价为 p^H，唐娜拒绝接受且还价为 p^L，那么______[填入接受或拒绝且以 p^M 还价]。

如果你列出所有的策略，你就会发现对马库斯来说，有 24 种不同的策略——换句话说，有 24 种填写空格的方法。

唐娜有四个信息集，她的策略模版如下：

如果马库斯要价 p^L，那么________[填入接受或拒绝]。

如果马库斯要价 p^M，那么________[填入接受或拒绝且还价为 p^L，或拒绝然后离开]。

如果马库斯要价 p^H，那么________[填入接受或拒绝且还价为 p^L，或拒绝且还价为 p^M，或拒绝然后离开]。

如果马库斯要价 p^H，我拒绝且以 p^L 还价，马库斯拒绝且以 p^M 还价，那么______[填入接受或拒绝]。

对唐娜来说，有 48 种可行的策略。她的策略确实有很多种，但记住一个策略的本质是：不论唐娜发现自己处在博弈的哪个阶段，她的策略都会指导她如何采取行动。

假设马库斯和唐娜选择以下策略组合。对马库斯来说：

在首节点，提价为 p^H。

如果我提价为 p^M，唐娜拒绝接受这个要价且还价为 p^L，那么我拒绝。

如果我提价为 p^H，唐娜拒绝接受这个要价且还价为 p^M，那么我接受。

如果我提价为 p^H，唐娜拒绝接受这个要价且还价为 p^L，那么我拒绝再还价为 p^M。

对唐娜来说：

如果马库斯要价为 p^L，那么接受。

如果马库斯要价为 p^M，那么接受。

如果马库斯要价为 p^H，我拒绝且以 p^L 还价。

如果马库斯要价为 p^H，我拒绝且以 p^L 还价，然后马库斯拒绝且还价为 p^M，那么接受。

我们利用这个策略组合来决定博弈在逻辑上的序列行动，从而得到相关收益值。在首节点，马库斯的策略规定他的提价为 p^H。根据唐娜的策略，她拒绝接受这个要价，且还价为 p^L。对于这个还价，马库斯的策略指导他拒绝接受且还价为 p^M。最后，唐娜的策略使她接受马库斯的还价 p^M。于是博弈的路径如下：马库斯提价为 p^H，唐娜拒绝接受这个要价，提出买价 p^L，马库斯拒绝且还价为 p^M，然后唐娜接受这个价钱。于是交易达成，车的卖价为 p^M。在这个策略组合中，唐娜得到的收益为 (p^M-p^M)，或为 0；马库斯得到的收益为 $2(p^M-p^L)$。

2.6 策略式转化为扩展式

尽管每个扩展式博弈只可以转化成一个相对应的策略式博弈，策略式却可以转化为多个扩展式博弈。这意味着在将扩展式转化为策略式时，我们丢失了一些信息，但这些丢失的信息无关紧要。

图 2—15 中描述的两个扩展式博弈，都可以转化为图 2—11 中的策略式博弈。在图 2—15a 的博弈中，斯卡比亚首先做出决定，然后托斯卡做决定，但托斯卡只有一个信息集，这表明当她决定刺杀或屈服时，她不知道斯卡比亚的决定是什么。这时，我们直接证明这个扩展式博弈可以转化成图 2—11 描述的策略式博弈。

除了参与者的序列行动不同，图 2—15b 中描述的博弈和图 2—15a 描述的博弈相同，它们可以转化成同一

个策略式博弈，得出这一结论有重大的意义。前面我们已经证明了博弈中重要的不是行动的顺序，而是当参与者行动时所知道的博弈信息。在这两个扩展式博弈中，当斯卡比亚采取行动时，他不知道托斯卡的选择：在图2—15a 中描述的博弈中，因为他比托斯卡先采取行动；在图 2—15b 中，因为他的信息集包括了托斯卡的两种行动方案。同理，在这两个博弈中，当托斯卡采取行动时，她也不知道斯卡比亚的选择。

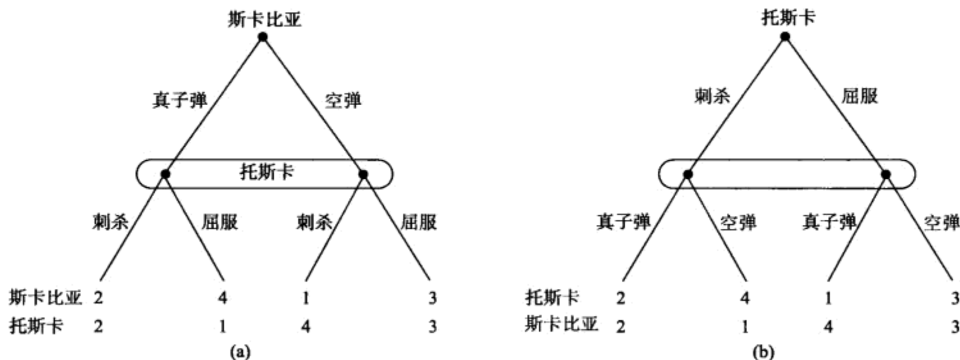


图2—15　两个扩展式博弈转化为同一个策略式博弈

2.7 共同知识

杰克和凯特准备在纽约的一家法国餐厅 PS 见面。杰克了解到这家餐厅今天停业，所以他发邮件给凯特说他们晚上 7 点整在另一家餐厅 AD 见面。凯特从她的黑莓手机收到邮件后回复邮件给杰克，说她 7 点准时到 AD 餐厅。杰克收到了她的回信。凯特 7 点准时出现在 AD 餐厅，但杰克不在那里。她就想杰克是否收到了她的回信。如果他没收到，那么他可能会认为她没有收到邮件，所以他可能去了 PS 餐厅。现在是 7 点 15 分，杰克仍然没到 AD 餐厅，所以凯特离开 AD 餐厅去了 PS。结果是杰克迟到了，当他到 AD 餐厅时发现凯特不在那里，觉得很奇怪。

杰克和凯特之间出现的问题就是博弈论中所讲的**共同知识**（common knowledge）的缺乏。杰克知道 Per Se 今天停业。凯特也知道 Per Se 今天停业，因为她收到杰克通知她的邮件。杰克知道凯特知道餐厅停业，因为他收到了凯特的回信。显然，凯特知道杰克也知道餐厅停业。但凯特不知道杰克知道她知道餐厅停业的消息。问题的关键是杰克和凯特都知道餐厅今天停业的消息是不够的：他们还需要知道对方知道的信息。

为了给共同知识一个更准确的定义，我们将一个事件（或一条信息）定义为共同知识，令 E 表示这个事件。在接下来的例子中，E 代表“PS 今天停业，在 AD 餐厅见面。”如果：

- 参与者 1 知道 E 且参与者 2 知道 E。
- 参与者 1 知道参与者 2 知道 E，且参与者 2 知道参与者 1 知道 E。
- 参与者 1 知道参与者 2 知道参与者 1 知道 E，且参与者 2 知道参与者 1 知道参与者 2 知道 E。
- 参与者 1 知道参与者 2 知道参与者 1 知道参与者 2 知道 E，且参与者 2 知道参与者 1 知道参与者 2 知道参与者 1 知道 E。
- ……

那么，对参与者 1 和参与者 2 来说，E 是共同知识。

我们列出所有的情况了吗？没有，因为这个逻辑可以无限推导下去。共同知识就像两个对立的镜子中无限反射的数字一样。这里，“反射”的是一个参与者知道的情况。那么共同知识包含的信息比参与者知道的信息多：它涉及其他参与者知道的信息，和其他参与者知道的其他参与者知道的信息等。

共同知识这个概念至关重要，因为在这本书提到的博弈中，潜在的假设是：**对每个参与者来说，博弈中的情形是共同知识**。每个参与者知道自己正在参与的博弈

是什么，每个参与者知道其他参与者知道自己正在参与的博弈是什么，如此等等。

当然，我们只是抽象地说明了什么是共同知识，那么任何事情（或信息）都是共同知识吗？即使我们紧挨着一起看电视，电视屏幕底边闪过天气预报，我能确定你看到了天气预报吗？可能可以确定。但我可以确定你看见我看到了天气预报吗？可能就不确定了。实际上，尽管知识可能包含不了需要满足共同知识条件的无限层面的信念，但它在许多集合中是最接近的一个信念，因为对参与者来说，有些事情“好像”是共同知识。

在我们分析之前，让我先描述一下和共同知识重要性有关的一个有趣的谜题。一群最近刚去世的人站在天堂的珍珠门前。圣彼得正在那里等着他们。他告诉这些人只有圣人才能进来，且圣人的头上有一轮光环。那些不是圣人但企图进入的将被送入地狱。那些不是圣人但也不企图进入的先送进炼狱，一会再送入天堂。确定你是不是圣人的关键是：没有人可以看见自己的头上有没有光环，尽管他/她可以看见其他人的头上有没有光环。圣彼得提供了最后一条信息：他宣布在这些人当中，至少有一个是圣人。

圣彼得开始邀请人们走进这扇门。如果没有人走进这扇门，那么他再次邀请，等等。这些人可以知道他们当中谁是圣人吗？如果“在他们当中至少有一个是圣人”是共同知识的话，那么他们可以知道当中谁是圣人。

为了更清楚地说明这个推论的过程，假设在他们当中只有一个圣人。刚开始这不是共同知识；每个人只知道在他们当中至少有一个是圣人。是圣人的那个环顾四周，看到其他人的头上都没有光环。因为他知道在他们当中至少有一个是圣人，他就可以得出这个结论：他头上一定有光环。于是他向天堂走去。你必须承认这个推论其实很简单，所需要知道的仅仅是“在他们当中至少有一个是圣人”。

现在我们假设他们当中有两个圣人，分别是蒂龙和丽塔。在圣彼得第一次邀请他们走进天堂时，每个人都环顾四周，看到至少有一个人头上有光环。（那些不是圣人的将会看到两个人头上有光环。）没有人可以确定他/她是圣人，所以没有人向天堂走去。记住：每个人都知道其他人也知道在他们当中至少有一个是圣人。没有人向天堂走去说明每个人一定看见了至少有一个人头上有光环，如果她没有看见，那么她就可以推断自己是圣人。因为每个人（也包括圣人），看见至少有一个人头上有光环，所以在他们当中一定至少有两个是圣人。既然蒂龙看到另一个人丽塔头上有光环，且他知道他们当中至少有两个圣人，蒂龙于是推断自己一定是圣人。同理，丽塔也可以推断出自己是圣人。在圣彼得第二次邀请他们时，蒂龙和丽塔同时向天堂走去。

好吧，我们已经解决了他们中有两个圣人的问题。但如果他们当中有三个圣人时，我们如何分析呢？大多数情况下，我们可以着手分析的条件是：每个人都知道他们当中至少有一个圣人，每个人都知道其他人知道他们当中至少有一个圣人，每个人都知道其他人知道其他人知道他们当中至少有一个圣人。你敢和我一同做这个吓人的逻辑推导吗？

假设他们当中有三个圣人，他们的名字分别是琼，米格尔和塔姆亚。与他们当中有两个圣人的情况一样，没有人可以推断出自己是圣人。因为每个人都看见至少有两个人头上有光环（琼、米格尔和塔姆亚都看见两个人头上有光环，其他人看见三个人头上有光环），知道“他们当中至少有一个是圣人”并不能推断出自己是圣人。所以，在圣彼得第一次邀请时，没有人走向天堂。因为没有人走向天堂，正如我们刚才推断的，每个人推断每个人都一定看见了一个圣人，即他们当中至少有两个圣人。所以圣彼得第二次邀请时的情况如何呢？因为实际上有三个圣人，在第二次邀请时每个人至少看见两个圣人。这种情况下，还是没有人能确定自己是圣人。现在，在圣彼得第二次邀请时，人们可以从剩余人群中推断出什么？每个人都推断出每个人（包括那些头上有光环的人）一定可以看见至少两个人头上有光环。所以，在他们当中至少有三个圣人。因为琼只看见两个人头上有光环（分别是米格尔和塔姆亚），琼头上一定有光环。所以在圣彼得第三次邀请时，琼、米格尔和塔姆亚都向天堂走去。通过三次邀请，这三个圣人确定了自己的身份。

假设他们当中有 n 个圣人，情况如何呢？同理可以推断出在第 n 次邀请之前，没有人向天堂走去。为了推导出这个结论，需要 n 层知识：每个人知道他们当中至少有一个圣人，每个人知道每个人知道他们当中至少有一个圣人，每个人知道每个人知道每个人知道他们当中至少有一个圣人，等等，直到第 n 个层面。于是，如果“他们当中至少有一个是圣人”是共同知识，那么其中

的圣人最后总能确认自己的身份。

博弈理论家认为要想成功，重要的不是你知道什么，而是你知道其他人知道的信息，及你知道其他人知道的其他人知道的信息，以及……

2.4 理解测验

写出图 2—8 中的抢劫博弈的策略式博弈。

2.8 模拟博弈时的其他相关问题

2.8.1　参与者还记得之前发生的事吗

在电影《记忆碎片》（*Memento*）中，主要人物莱昂纳多·谢尔比（Leonard Shelby）因为妻子被杀而心灵受到创伤，这使得他无法拥有长期记忆，他只能记住几分钟之前发生的事情。尽管记忆支离破碎，他仍发誓要追查到凶手。每当他找到一些有价值的线索时，他必须使用一切方法立即将之记录下来，以便过后再利用这些线索。对他来说最重要的线索是他身上的纹身。他会忘记所有没有被记录下来的信息。莱昂纳多认识到自己记忆的断裂，但事实上他策略性地利用了这个缺陷。有时候，他写下某些人的名字，认为这个人就是凶手，即使他知道这个人并不是凶手。但他也知道几分钟过后他还会看这些记录且相信他记录下来的信息，因为他忘记了这些都不是真的。通过这种方式，不久他就认为自己已经找到了凶手，追杀凶手的过程满足了他复仇的心理。

按照博弈论的说法，莱昂纳多·谢尔比拥有**不完美记忆**（imperfect recall），因为他现在不知道他之前知道的事实。尽管博弈论也承认这种记忆的存在，但这本书中我们主要研究**“完美记忆”**（perfect recall），所以博弈中的参与者知道博弈任何阶段中发生过的事实。比如，在伊拉克拥有大规模杀伤性武器博弈中，当伊拉克决定是否拒绝接受检查时（当联合国要求检查伊拉克是否拥有大规模杀伤性武器时），它记得自己是否拥有大规模杀伤性武器（博弈开始时的行动）。同理，在和汽车经销商讨价还价博弈中，如果马库斯刚开始要价 p^H，唐娜拒绝接受这个价且还价 p^L，当马库斯决定是否接受唐娜的还价时，他记得自己刚开始的要价 p^H 并第二次提出要价 p^M。在第 1 章提到的合并的博弈中，你可以承认不完美记忆，因为事实上，参与者的不完美记忆是那个博弈最核心的要素。而在本书中，不完美记忆不是我们的研究对象。

2.8.2　一个参与者可以改变博弈的进程吗

思考电影《赎金风暴》（*Ransom*）中的绑架事件。梅尔·吉布森（Mel Gibson）饰演主要人物约翰·马伦（Tom Mullen）。一个名为吉米的腐败警长绑架了约翰的儿子肖恩。一开始，这是一个标准的绑架博弈。吉米要求 200 万美元的赎金。在决定不交付赎金时，约翰就认定绑架者无意释放他的儿子。于是他决定去当地电视台播放一段视频：

> 现在全世界都知道我的儿子肖恩在三天前被绑架了，这是他最近的一张照片。肖恩，如果你正在看电视，我想告诉你，我们爱你。绑架肖恩的人，你看，这是你要求的赎金，200 万美元。但你一分钱也得不到，因为我不会交赎金，哪怕是一分钱。相反，我会把赎金奖赏给那些提供线索捉拿你的人。所以，恭喜你！就在刚才，你成为了 200 万美元彩票。你知道有人会为了 200 万美元把你交给警察吧？我觉得你不会这样想。但这是你的最后机会了，如果你放了我的儿子，没让他受到任何伤害的话，我会立即撤销奖金。这样，你也可以幸运地逃脱。理解了吗？你永远不会得到这笔钱！所以你还有改邪归正的机会。如果你执意要犯错，那么愿上帝保佑你！因为在这个世界上不会有人保佑你了！

很显然，约翰“灵机一动”把赎金转化成彩票。刚开始的博弈如图 2—1 中所描述的情形。但不久约翰把博弈“改变”了。遗憾的是，博弈论并不承认这些改变或创新。**博弈论的一个关键假设是参与者都知道博弈的状况；博弈的规则是共同知识。**

然而，我们可以做的是把博弈修正成图 2—16 中描述的博弈。现在约翰可以选择交付赎金、悬赏缉拿绑架者或者什么也不做。当开始博弈时，所有参与者都知道博弈的情形，所以博弈不允许约翰“改变”博弈，把赎金变为赏金。这个创新不是博弈论的特征，所以，参与者不可以改变博弈的进程。然而，在博弈中，我们可以

给参与者更多的选择。

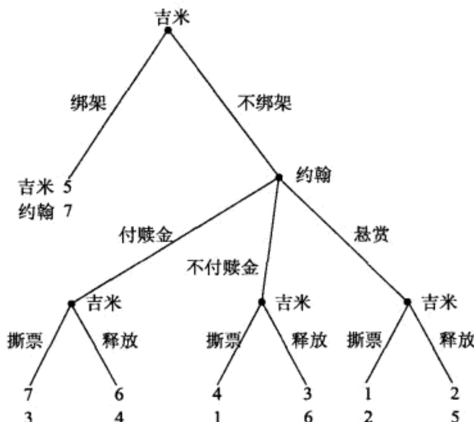


图 2—16　电影《赎金风暴》中的扩展式博弈

2.8.3　博弈中的事实一定准确吗

如果我们构建和求解博弈的目的是为了理解参与者的行为，那么最关键的不是实际上事实的真假，而是参与者所认定的事实的真假 。参与者的行为受到他们的偏好和信念的驱使，不论这些事实是真还是假。所以，一个博弈应当描述的是参与者所认定的博弈形势。

如果一个参与者是 19 世纪印第安部落中的一员，他认为部落首领有神奇的力量。要理解他们的行为，我们就需要承认这个信念——不论这个信念是真还是假。或者，如果一个生活在 21 世纪的参与者相信将飞机撞向建筑物可以谋得来世的福祉，同理，我们需要承认这个信念，不论这个信念多么可笑，多么荒唐。

本章总结与练习题

本章总结

当法国印象派画家克劳德·莫奈（Claude Monet）观看伦敦的一个建筑物、法国的一座教堂，或者是他家后院的一个荷花池时，他依靠对这些物体的印象而不是靠实物完成画作。模拟现实生活中人们的交易就像模拟一件艺术品。现实生活是复杂凌乱的，如果一个社会科学家要构建一个简单的模型来理解生活中的一切，他就必须从模型中提炼出博弈的本质特征。要提炼出这些本质特征，一个社会科学家就必须具有创造力、洞察力和判断力。博弈论为那些具备这些能力的智者提供了构建模型的工具。

在这章，我们回顾了构建策略情境中博弈理论模型的两种方法。一个是扩展式博弈，它借用树型结构来描述参与者采取的序列行动并描述了这些行动的博弈环境，包括可供参与者选择的行动方案与博弈中参与者所面临的博弈状况。这些知识表现为**信息集**，信息集是指对于特定的参与者，他所观察到的所有博弈中可能发生的行动的集合。信息集使我们可以构建不同形势下的博弈模型，即使在构建模型时，我们缺乏相关信息。一个信息集既可以体现**完美信息博弈**（其中一个参与者知道博弈中各个阶段的状况）的情况，也可以体现**不完美信息博弈**（其中一个参与者不确定其他参与者将如何采取行动）的情况。描述参与者行为的关键是了解参与者看重的是什么，所以一个**扩展式博弈**也可以描述参与者的**收益**（一个参与者赋予博弈结果的数值）。

策略式博弈的描述方式比扩展式更简洁。策略式博弈的三要素是：**参与者集合、每个参与者的策略集和每个参与者的收益函数**。参与者的决定是从他/她的策略集选择的行动方案，这里策略是决策规则——指导参与者选择行动方案。收益函数告诉我们一个参与者如何评估任何一个策略组合（在博弈中，每个参与者有一个策略集）。

正如在随后的几章中所讲述的那样，对预测和描述参与者的行为最重要的是了解每个参与者知道其他

参与者知道的知识。这个知识包括每个参与者对其他参与者的信念。一个最重要的假设是：对参与者来说，“博弈中的情况”是**共同知识**。这不仅表明参与者同意博弈规则，还表明每个参与者知道其他参与者知道的博弈的状况，等等。共同知识就像一个完美的圆，在现实中并不存在这样一个概念，但无论如何这是我们研究人类所生活的世界的一个有用的抽象概念。

练习题

1. 大洋洲和欧亚大陆的国家正在交战。如图 2—17 中所描述的博弈，大洋洲有四个城市：A，B，C，和 D，相传欧亚大陆要轰炸这四个城市中的一个。轰炸机可能来自 X 基地（可以到达 A 和 B 这两个城市），也可能来自 Y 基地（可以到达 C 和 D 这两个城市）。欧亚大陆决定攻打四个城市中的一个。大洋洲不知道欧亚大陆选择的是哪个城市，但可以观测到基地上空的轰炸机。在观测后，大洋洲决定撤离其中一个城市（仅撤离一个城市）的市民。假设如果大洋洲成功撤离了被轰炸城市的市民，得到的收益为 2，如果撤离的是其他三个城市中的任意一个，得到的收益为 1。假设欧亚大陆轰炸的是市民未撤离的城市，得到的收益为 1，如果轰炸的是成功撤离了市民的那个城市，得到的收益为 0。列出扩展式博弈。

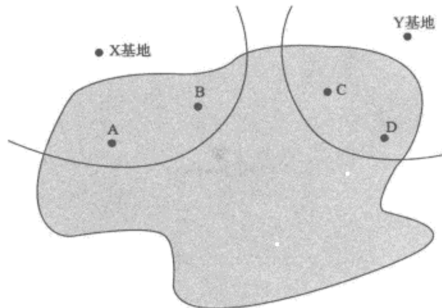


图 2—17

2. 参与者 1 先行采取行动，可供选择的行动方案有：a，b，c 和 d。如果参与者不选择 d，那么参与者选择 x 或者 y。参与者 2 可以观测到参与者 1 的行动。如果参与者 1 选择 d，那么参与者 3 选择左或者右。列出这个博弈的扩展式，可以不列出得到的收益值。

3. 思考在一个博弈中，参与者 1 先行采取行动，他的可行选择有：a，b 和 c。在观测到参与者 1 的行动后，参与者 2 选择行动 x 或者 y。思考以下三种情况下参与者 3 可以采取什么行动以及当她采取行动时她了解到的信息：

a：如果参与者 1 选择 a，那么参与者 3 在行动高和低中选择。当参与者 3 采取行动时，她知道参与者 2 的选择。列出这个博弈的扩展式，可以不列出得到的收益值。

b：如果参与者 1 选择 a，那么参与者 3 在行动高和低中选择。当参与者 3 采取行动时，她不知道参与者 2 的选择。列出这个博弈的扩展式，可以不列出得到的收益值。

c：如果参与者 1 选择 a 或者选择 b，那么参与者 3 在行动高和低中选择。参与者 3 观测到参与者 2 的选择，但不知道参与者 1 的选择。列出这个博弈的扩展式，可以不列出得到的收益值。

4. 思考 2.1 小节中的美国巡回上诉法院的博弈。

a：假设在博弈开始时，所有的陪审员都知道陪审员 Z 只能阅读海瑟芙呈交的报告。列出这个扩展式博弈，可以不列出得到的收益值。

b：假设在博弈开始时，所有的陪审员都知道陪审员 X 先行投票，且当陪审员 Y 和陪审员 Z 同时投票时，他们都观测到了陪审员 X 的投票结果。列出这个扩展式博弈，可以不列出得到的收益值。

5. 市议会正举行有关提高物业税的听证会。假设 T 女士是听证会的主席，议会的其他两名成员分别是 T 先生和 D 夫人。投票程序如下：除了主席，T 先生和 D 夫人同时在纸上写下他们的投票结果。他们或者支持提高物业税，或者反对提高物业税。然后市议会的秘书打开纸条，公开宣布投票结果。如果这个秘书报告两个纸条上都是支持，那么实施提高物业税，博弈结束。如果两个纸条上都是反对，那么不实施提高物业税，博弈结束。然而，如果一个支持，一个反对，那么 T 女士必须投票。如果她投票支持，那么实施提高物业税，博弈结束；如果她投票反对，那么不实施提高物业税，博弈结束。至于收益情况，如果实施提高物业税，那么 T 女士，J 先

生和D夫人每个人得到的收益为3。如果提高物业税的提议失败，那么D夫人得到的收益为4，J先生得到的收益为1。对T女士来说，她希望提高物业税——相信提高物业税将会为提高学校的质量提供资金，但如果提议是将物业税提高到一个更高水平，那么她反对这项提议。当J先生和D夫人都支持这个提议时，她得到的收益为5；当J先生和D夫人中有一个人支持，一个人反对且她自己投票支持时，她得到的收益为2。如果不实施提高物业税，那么她得到的收益为0（不论她是否投票）。列出这个博弈扩展式。

6. 思考在一个传奇的博弈秀节目《来做个交易吧》（*Let's make a deal*）中的一个选手面对的情况。一共有三扇门，其中两扇门后是傻子奖（也就是奖金价值最小的奖项），一扇门后是丰厚的大奖，比如一辆汽车。这些门分别被编号为1、2和3。在节目开播之前，主持人选择在一扇门后放置大奖。然后，在节目中，这个选手选择三扇门中的一扇。之后，主持人选择打开没被选手选中的那两扇门的其中一扇。这个节目的一个游戏规则就是主持人不能打开那扇有大奖的门。在主持人打开门之后，再给选手一次机会继续选择，他可以继续选择第一次选择的那扇门，也可以选择那扇没被打开的门。在选手做出选择后，打开剩下没被打开的两扇门。

a：列出这个博弈的扩展式。包括的博弈阶段是：从博弈开始到选手决定是选择第一次选择的那扇门还是选择那个没被打开的门这个阶段。所以，你列出的扩展式博弈应该包括以下三个阶段：（1）主持人选择一扇门放置大奖，（2）选手选择一扇门，（3）主持人选择打开一扇门（可以不列出得到的收益值）。

b：在选手决定是继续选择第一次选择的那扇门还是选择没被打开的另一扇门这个阶段中，写出选手的信息集。在节点旁标上数字来描述通向那个节点的序列行动，如3/2/1。3/2/1表示主持人将大奖放在3号门后，选手刚开始选择2号门，主持人打开了1号门。

7. 返回到图2—10中描述的伊拉克战争博弈，写出三个参与者的策略集。

8. 看图2—18中描述的扩展式博弈，推导相对应的策略式。

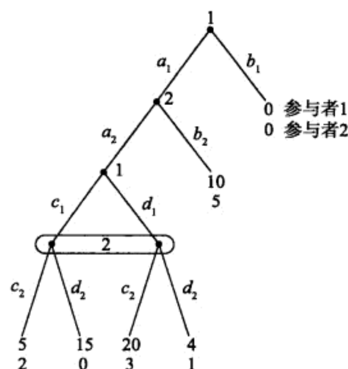


图2—18

9. 列出图2—19中描述的扩展式博弈相对应的策略式博弈。

10. 列出图2—20中描述的扩展式博弈相对应的策略式博弈。

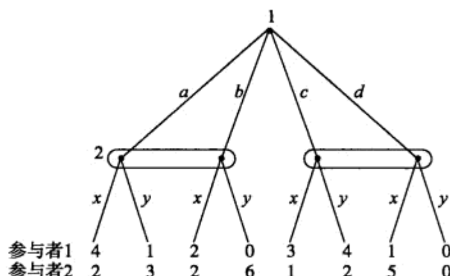


图2—19

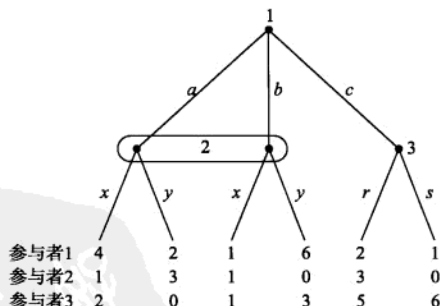


图2—20

11. 看图 2—21 中描述的三个扩展式博弈。如果它们当中有违反完美记忆的假设，陈述这些违反完美记忆的博弈并给出你的解释。

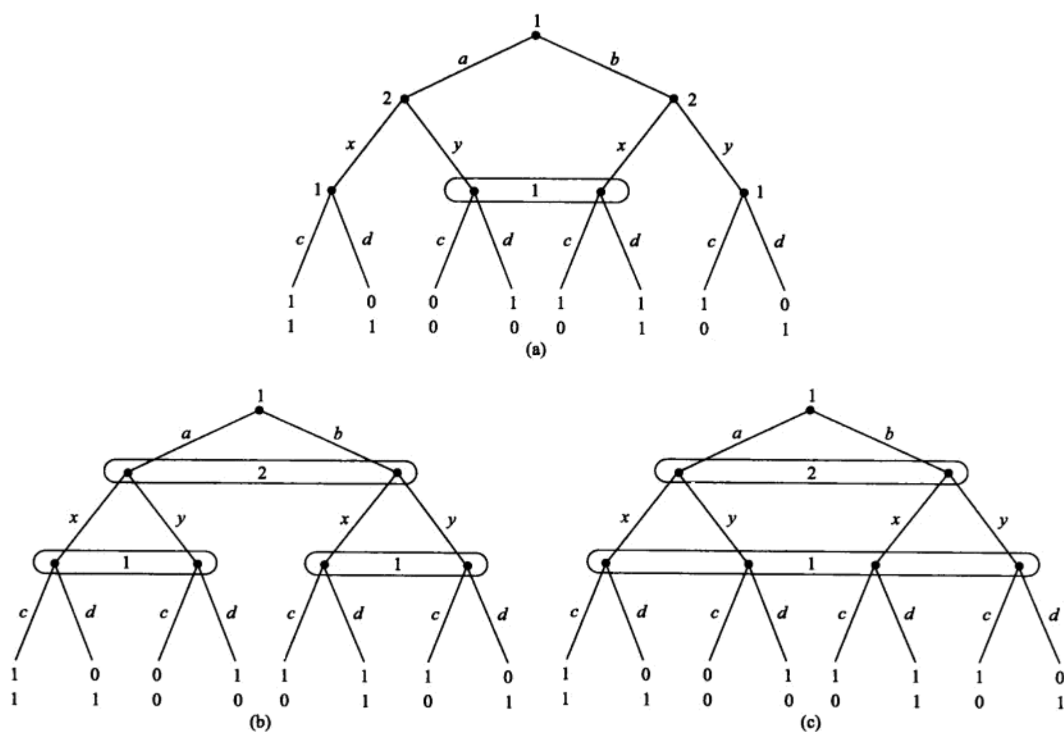


图 2—21

PDG

PDG

第二部分　求解策略式博弈

PDG

第 3 章

剔除不可能：理性是共同知识下的博弈求解

学习目标

3.1 博弈求解：参与者是理性的

3.2 博弈求解：参与者是理性的和参与者知道参与者是理性的

3.3 博弈求解：理性成为共同知识

本章导读

> 我不是曾经和你说过多次吗？当你把不可能的因素都剔除出去以后，不管剩下的是什么，不管是多么令人难以置信的事——都是实情！
>
> ——福尔摩斯，《血字的研究》

在第 2 章中，我们了解到如何构建扩展式和策略式的博弈模型。但是，在构建起一种博弈模型之后，我们用它来做什么？在回答之前，我们先来扮演一回业余侦探，研究一下人们在策略情境中的行为反应。要做到这一点，我们需要知道如何求出博弈的解，这也是本章的核心所在。

图 3—1 所示的是策略式绑架博弈（其扩展式见图 2—6）。回忆一下那名绑架者盖伊，他有四种策略：（1）不绑架奥兰多，但事实上他不仅这么做了，而且还杀了人质（不绑架 / 杀人）；（2）不绑架奥兰多，但事实上他这么做了之后，又放了人质（不绑架 / 释放）；（3）绑架奥兰多而且杀了他（绑架 / 杀人）；（4）绑架奥兰多然后释放他（绑架 / 释放）。奥兰多的家属薇薇卡可以选择支付赎金或者不支付赎金。单元格中的第一个数字代表横排参与者的收益（盖伊的收益），第二个数字表示的是竖行参与者的收益（薇薇卡的收益）。注意，收益数字是按从最差结果到最佳结果的顺序排列的。

会发生绑架吗？如果发生绑架，会支付赎金吗？奥兰多能够生还吗？要求出这一博弈的解，就意味着要从这八对可能的策略中作出选择来回答这些问题。我们需要剔除不合理和不可行的策略组合，然后从理论上找出一个独特而强有力的策略组合。解决方案越少，我们对行为作出的预测就越精确。要推导出一个解决方法，就需要假定某一参与者在众多的策略中是如何作出选择的。当然，这项工作的挑战性就在于一个参与者选择策略的方式很可能取决于她对其他参与者如何做选择的判断。这是一个不容易解决的复杂工程，那么，我们最好以简单的方式着手进行。

PDG

		薇薇卡（受害者家属）	
		支付赎金	不支付赎金
盖伊（绑架者）	不绑架/撕票	3,5	3,5
	不绑架/释放	3,5	3,5
	绑架/撕票	4,1	2,2
	绑架/释放	5,3	1,4

图 3—1　策略式绑架博弈

我们计划以循序渐进的方式对参与者做出更多的假设，随后挖掘他们行为背后的原因。起始阶段，我们会假设参与者是理性的（见 3.1 小节）；接着，会进一步假设每一个参与者相信所有的参与者都是理性的（见 3.2 小节）；然后，在此基础之上，假设每一个参与者相信所有的参与者都相信所有的参与者都是理性的（见 3.2 小节）；最后，在 3.3 小节中总结求解技巧后结束本章内容。

3.1 博弈求解：参与者是理性的

要构建某一参与者策略选择的模式，需要先假设所有的参与者都是**理性的**（*rational*）。当一位参与者为尽可能极大化自己的利益而采取行动时就被认为是理性的。或者更确切地说，根据对其他参与者会采取何种行为的信念，某一参与者会采取使自己的收益最大化的策略。需要注意的是，这里所说的理性并不是指对于其他参与者行为应具有怎样合理的判断，而是指一位参与者根据其他参与者作出何种策略的信念，选择最大化自己收益的策略。

人们只以自身的利益作为行动的出发点合理吗？这种假设是否意味着人们都是自私自利的？虽然理性确实包含追求自身利益，但是却没有明确规定是什么利益。它或者是指伊本内塞•史克鲁奇（Ebenezer Scrooge）圣诞前夜满是金钱的想法，或者是指那顿悟之夜过后他对所有同胞的关爱。理性就像是一个向所有人都敞开大门的教堂，这其中既有利己主义者也有利他主义者。理性意味着仅仅追求自身的利益，而不管这些利益是如何被定义的。

最初，我们先假定理性具有某一特定的含义，即不论其参与者采取什么样的策略，只要另外一个策略（称之为 s''）总是能够产生一个严格更高收益，那么这位参与者就不会采用现有的某一策略（称之为 s'）。因此，s' 策略永远不是最佳策略，因为 s'' 总是会产生更大的收益。所以，不管你是特蕾莎修女（Mother Teresa）还是帕里斯•希尔顿（Paris Hilton），选择策略 s' 都是相当愚蠢的行为。我们进一步深入，假设每一位参与者相信其他的参与者相信他们不会选择愚蠢的策略，以此类推。换句话说，参与者的共同知识是如果另一个策略产生严格更高收益，那么参与者就不会选择现有的某一特定策略。

在附录 1 中，我们介绍了**可理性化**（rationalizability）这样一个概念，这是理性是共同知识内涵的进一步深化。从本质上来说，可理性化更接近于理性。正是因为可理性化是一个相当微妙而又复杂的概念，所以我们在这里把它作为一个选读内容，供那些更富有冒险精神的学生或者是更加严格的教师涉猎和研究。

3.1.1　严格占优

回到意大利歌剧作家普契尼的歌剧《托斯卡》中，图 3—2 重现了这一策略式博弈。在 2.4 小节中，斯卡比亚，那名警察局长，判处弗洛里亚•托斯卡的恋人马里奥•卡瓦拉多西死刑。斯卡比亚告诉托斯卡如果他能够得到她，就会让执行死刑的射击队发空弹，让卡瓦拉多西不死。斯卡比亚首先命令射击队发真子弹或是空弹，然后去见托斯卡，此时，托斯卡必须决定是屈服于斯卡比亚还是刺杀他。紧接着，卡瓦拉多西被带到了射击队前执行枪决。托斯卡和斯卡比亚在互相不知道对方选择的情况下各自采取行动。收益显示出托斯卡最关心她的恋人是否能得救，而后才是不屈服于斯卡比亚；而对于斯卡比亚来说，他渴望能够拥有托斯卡，卡瓦拉多西是否执行枪决倒在其次。

		斯卡比亚	
		真子弹	空弹
托斯卡	刺杀	2,2	4,1
	屈服	1,4	3,3

图 3—2　托斯卡博弈

如果我们假设这两位歌剧人物是理性的，就能够预测他们的行为吗？假如斯卡比亚选择发真子弹，那么托斯卡选择刺杀的收益是 2，选择屈服的收益是 1（见图 3—3）。所以，当托斯卡认为斯卡比亚下令射击队发真子弹时，显而易见，她会选择刺杀他。而如果她认为斯卡比亚会命令发空弹，情形又会是怎样的呢？选择刺杀的收益是 4，而选择屈服的收益是 3，因此，她会再一次选择刺杀。那么，不论托斯卡认为斯卡比亚会怎么做，选择刺杀的策略都比选择屈服产生严格更高收益。所以，托斯卡肯定不会选择屈服。既然必须采取行动，那么对于托斯卡而言，只有刺杀斯卡比亚。从另外一个角度来说，不管斯卡比亚采取什么行为，刺杀的策略都远远优于屈服，所以一个理性的托斯卡很可能会选择刺杀的策略。

		斯卡比亚	
		真子弹	空弹
托斯卡	刺杀	2,2	4,1
	屈服	1,4	3,3

图 3—3　托斯卡在斯卡比亚选择下令发真子弹情形下的收益

同理，斯卡比亚的理性会促使他选择下令发真子弹的策略。如果托斯卡选择刺杀他，他选择杀掉卡瓦拉多西的收益是 2，而不这么做的收益则是 1；假如托斯卡选择屈服，那么，他选择发真子弹的收益是 4，又一次超过了选择发空弹的收益 3。因此，无论斯卡比亚认为托斯卡采取什么行动，他都会下令行刑队发真子弹。总而言之，博弈论可以做出清晰的预测（虽然这样的预测带有一些血腥的色彩），即托斯卡会刺杀斯卡比亚，而斯卡比亚会让卡瓦拉多西死在行刑队的手中。

对于托斯卡而言，屈服策略和刺杀策略相比属于**严格劣势策略**（strictly dominated）。也就是说，不管斯卡比亚选择何种策略，托斯卡选择刺杀带来的收益要高于选择屈服。

定义 3.1

严格优于（strictly dominates）：无论其他参与者的策略是什么，若策略 s' 产生的收益严格高于 s'' 产生的收益，那么策略 s' 为严格优于策略 s''。[1]

一个严格优于其他策略的策略对于参与者而言就是**占优策略**（dominant strategy）。显然，在只有两种策略的情况下，如果托斯卡的屈服策略属于严格劣势策略，那么，刺杀策略则属于占优策略。

定义 3.2

占优策略（dominant strategy）：如果一个策略严格优于其他策略，那么称这个策略为占优策略。

如果一个策略是严格劣势策略，那么关于其他参与者采取何种策略的任何信念都不是最佳策略。因而，理性的参与者都不会选择这样的策略。此外，如果一个参与者有一个占优策略，而且是理性的，那么他就会使用这一策略。每一个参与者都拥有一个占优策略时，唯一的解就是每一个参与者运用属于自己的占优策略。



点拨

一个理性参与者从不使用严格劣势策略，一个理性参与者总是使用占优策略。

在将这些新的方法应用到其他博弈情形之前，先来关注一下歌剧《托斯卡》结局所具有的一个颇为有趣的特性。托斯卡选择刺杀斯卡比亚，斯卡比亚选择让行刑队枪毙卡瓦拉多西时，他们各自的收益是 2。而如果他们选择另一策略组合，即托斯卡选择屈服，斯卡比亚下令让行刑队发空弹，那么卡瓦拉多西就可以得救。托斯卡和斯卡比亚各自得到的收益会是 3，而且各自的处境会好很多。

歌剧《托斯卡》的博弈情境很好地说明了个人理性和集体理性的重要区别。无论是托斯卡选择刺杀斯卡比

①定义 3.1 和 3.2 更为正式的数学定义参见本章附录。——作者注

亚（这是她的占优策略）还是斯卡比亚选择命令行刑队发真子弹（这是他的占优策略），都是个人理性的表现。然而，集体理性则意味着每一个人的处境都会有所改善，表现为托斯卡和斯卡比亚分别选择屈服和发空弹的策略。因而，个人利益最大化不一定意味着集体利益最大化。关于这一点我们会在后面的章节做进一步的研究。

那么，该歌剧中人物的结局是怎样的呢？托斯卡刺死了斯卡比亚，斯卡比亚下令行刑队发真子弹，结果卡瓦拉多西和斯卡比亚都死了。当托斯卡得知卡瓦拉多西的死讯时，她从城堡的壁垒纵身跳下结束了自己的生命。尽管充满了残杀，但爱终究战胜了色欲。或者，假如这是一场网球比赛，那么爱与色欲的比分是 15 比 0。



托斯卡刺杀斯卡比亚场景

案例研究

白人迁徙与租住房中的种族隔离

你认为搬到一个自己并不受欢迎的街区会给你带来什么好处，尤其是当那里的人们感到整个生活方式和自己为之奋斗的一切都受到了威胁而变得非常焦躁不安的情况下。①

在 20 世纪 60 年代，把白人街区的房子或公寓出租或出售给美国黑人是一个颇受争议的种族问题。"白人迁徙"一词是指随着一些黑人家庭的到来，白人家庭开始大批迁移出去，使得街区由"白人化"转变为"黑人化"。如果白人和黑人愿意居住在种族隔离的街区，会发生白人迁徙吗？

设想一个黑人家庭愿意支付比白人家庭更高的租金租住在一个白人街区中。支付更高租金的意愿就表明了对于黑人家庭来说几乎没有机会选择漂亮的房子和良好的学校。假设在某一个街区有八座类似的房子，表 3—1 列出了根据居住在该街区黑人家庭的数量，假定一个黑人家庭和一个白人家庭愿意支付的月租金。注意黑人家庭和白人家庭都愿意在一个有些种族隔离的街区并支付最高租金。

表 3—1　住宅区市场的租金

黑人家庭数量	黑人家庭支付的租金	白人家庭支付的租金	合计
0	–	100	800
1	110	105	845
2	115	110	890
3	120	100	860
4	110	90	800
5	100	75	725
6	90	75	690
7	85	70	665
8	80	–	640

合计一栏是指房主从八座房子获得的月租金总和。假如有 3 个黑人家庭，每个黑人家庭月租金 120 美元，每个白人家庭月租金 100 美元，总计 860 美元。仔细观察，当其中两座房子租给黑人家庭时，房东月租金总额最高可达 890 美元。

博弈的焦点并不在黑人和白人家庭之间，而是在八座房子的房主之间。假使每一座房子的房主都不同并且每一个房主的收益等同于其所收的租金，那么在决定由谁租住自己的房子时，房主选择的策略会不会产生一个种族隔离的街区？又或者房主们的决定会使得该街区由"白人化"转向"黑人化"吗？

严格占优的概念会帮助我们预测出事情的结果。先考虑在其他七个房主都将房子租住给白人家庭的条件下，一个房主的决策。如果第一个房主把房子租给一个黑人家庭一个月会赚 10 美元收益（110 美元减去 100 美元），那么，当其他七位房主把房子都租给白人家庭时，他愿意把房子租给一个黑人家庭。相反，假如其他七座房子中有六座租给白人家庭，一座租给了一个黑人家庭，

① （白人）卡尔·德林在面对莉娜·杨格（黑人）购买了自己所住街区的一幢房子时，对杨格一家说的话，出自《日光下的葡萄干》（*Aratsin in the Sun*），洛兰·汉斯贝里（Loraine Hansberry）。——作者注

则如果上述房主将房子租给一个黑人家庭，他的收益将会是 115 美元（因为目前街区有两个黑人家庭），超过了租给一个白人家庭 105 美元的收益（当时只有一个黑人家庭）。因而，在其他七座房子中有六座租给白人家庭的前提下，第一位房主把房子租给一个黑人家庭会对其有利。以此类推，你会发现无论其他七位房主把房子租给谁（就房客的种族而言），一房主把房子租给一个黑人家庭会获利更多。也就是说，把房子租给一个黑人家庭的策略严格优于租给一个白人家庭。

禁止电视播放香烟广告

尝过了云斯顿（Winston）香烟才知道香烟应有的味道！①

生活在 20 世纪 60 年代的人们都不会对电视上的云斯顿香烟广告的流行用语感到陌生。事实上,尽管它已经传播了 30 多年，但包括我在内的许多人对其仍是记忆犹新。1971 年联邦政府颁布了香烟广告禁令后，美国的广播电视中就再也看不到（听不到）香烟制品的广告了。

通常，政府实施某一限制是公司不希望看到的，因为剥夺了选择权就意味着限制了获取高额利润的渠道。然而，博弈理论认为失去选择权并非总是坏事。说不定正是禁止电视广播播放香烟广告的做法增加了香烟生产者的利润。

以菲利普莫里斯公司（Philip Morris）和雷诺烟草公司（R. J. Reynolds）为例，二者一直都是最大的香烟生产商。多数人都熟悉他们的万宝路和云斯顿香烟品牌。了解广告是如何影响香烟的销售量对于探讨广告的需求量具有十分重要的作用。经济学家的一些研究表明，广告对吸烟人群的数量并没有太大的影响，更多的是将现有的吸烟人士分流到不同的品牌。但是，也有其他的证据表明香烟广告不仅可以维持现有吸烟人群，而且还可以吸引不吸烟者（特别是年轻人）尝试吸烟。为了简化我们的研究模式，姑且认为香烟广告不会影响香烟的销售总量，而只是引导不同香烟品牌的受众。

假定香烟的年需求量是 1 000 000 000（十亿）包，每家公司的市场份额由其自身与其对手制定的广告费决定。ADV_{PM} 指菲利普莫里斯公司的广告费，ADV_{RJR} 则表示雷诺烟草公司的广告费用。假设菲利普莫里斯公司的市场份额等于：

$$\frac{ADV_{PM}}{ADV_{PM}+ADV_{RJR}}$$

这个公式表示出菲利普莫里斯公司的市场份额等同于它的广告份额。虽然这样的研究模式有些过于简单，但我们真正需要假定的是广告费用越高香烟的销售量越大。因而，菲利普莫里斯公司的销售量等于：

$$1\ 000\ 000\ 000\times\left(\frac{ADV_{PM}}{ADV_{PM}+ADV_{RJR}}\right)$$

类似的，雷诺烟草公司的销售量是：

$$1\ 000\ 000\ 000\times\left(\frac{ADV_{RJR}}{ADV_{PM}+ADV_{RJR}}\right)$$

如果每包香烟获利 10 美分（记住我们要回到 1971 年），那么，菲利普莫里斯公司支出的广告费用产生的收益是：

$$0.1\times1\ 000\ 000\ 000\times\left(\frac{ADV_{PM}}{ADV_{PM}+ADV_{RJR}}\right)-ADV_{PM}$$

或：

$$1\ 000\ 000\ 000\times\left(\frac{ADV_{PM}}{ADV_{PM}+ADV_{RJR}}\right)-ADV_{PM}$$

同理，雷诺烟草公司的收益是：

$$1\ 000\ 000\ 000\times\left(\frac{ADV_{RJR}}{ADV_{PM}+ADV_{RJR}}\right)-ADV_{RJR}$$

我们的目标是计算出这些公司的广告费用。为了方便计算，假设只有三种等级的广告：500 万美元，1 000 万美元和 1 500 万美元。这样产生的收益矩阵如图 3—4 所示（当中策略和收益的单位都是百万美元）。

		雷诺烟草公司		
		花费5	花费10	花费15
菲利普莫里斯公司	花费5	45,45	28,57	20,60
	花费10	57,28	40,40	30,45
	花费15	60,20	45,30	35,35

图 3—4　香烟广告博弈

例如，假设菲利普莫里斯公司的广告花销是 500 万美元，雷诺烟草公司的广告花销是 1 500 万美元，则菲利普莫里斯公司的市场份额为 5/（5+15），或 25%，它的香烟销售量是 2 500 万包（0.25 乘以 10 亿），以每包 10 美分的利润计算，该公司总获利 2 500 万美元。剔除

①为了应对语法专家的批评，云斯顿提出了新的标语：“选择纯正语法还是纯正味道？”——作者注

广告成本费后，菲利普莫里斯公司的净利润（或收益）达到2 000万美元。



云斯顿香烟广告

表3—2说明1 500万美元广告花销策略严格优于500万美元或1 000万美元广告花销策略。比如，假如雷诺烟草公司的广告花销是500万美元，那么菲利普莫里斯公司1 500万美元的广告花销可以带来6 000万美元的利润（并且拥有75%的市场份额）；而1 000万美元的广告费用则产生5 700万美元的利润；但如果像雷诺烟草公司一样只花费500万美元打广告，利润则只能维持在4 500万美元。同样，面对雷诺烟草公司1 000万美元的广告花销，菲利普莫里斯公司1 500万美元的广告预算要远远优于其他两种选择。因此，不管雷诺烟草公司做出何种策略，菲利普莫里斯公司都会选择花重金进行广告宣传，大规模的广告宣传成为该公司的占优策略。雷诺烟草公司的情形也是如此，所以，可以料想两家烟草公司会通过电视节目中俊男美女吞烟吐雾的广告形象来推销自己的产品。

表3—2　1 500万美元的广告预算成为菲利普莫里斯公司的占优策略　单位：万美元

雷诺烟草公司备选策略	菲利普莫里斯公司花费1 500的收益		菲利普莫里斯公司花费1 000的收益		菲利普莫里斯公司花费500的收益
花费500	6 000	>	5 700	>	4 500
花费1 000	4 500	>	4 000	>	2 800
花费1500	3 500	>	3 000	>	2 000

假设现在禁止电视广播播放香烟广告，烟草公司无法运作1 500万美元的广告计划。也就是说，一个烟草公司的广告预算最高上限为1 000万美元。在这种简单博弈的情境中，每一个烟草公司只能在500万美元和1 000万美元之间选择自己的策略组合（如图3—5所示）。

		雷诺烟草公司	
		花费500	花费1 000
菲利普莫里斯公司	花费500	4 500, 4 500	2 800, 5 700
	花费1000	5 700, 2 800	4 000, 4 000

图3—5　禁止电视广播发布香烟广告情形下的香烟广告博弈

此项博弈的解就是两家烟草公司适度规划广告预算，因为1 000万美元花销要严格优于500万美元花销。那么政府这种倾入的政策会对他们的利润产生怎样的影响？答案是会增加他们的利润，即每家公司的利润会从3 500万美元上升至4 000万美元。

在原有的博弈中，每家公司都有一个1 500万美元花销的占优策略。这种重金打广告的做法被取消之后，每家公司的市场份额只剩下50%。如果他们将广告成本减少至1 000万美元，那么每家公司仍然占有一半的市场份额，因此可以保持总利润不变，而削减的广告费用转化为更高的净利润。

禁止电视广播发布烟草广播的政策，通过限制广告权限，实现了减少竞争、削减广告支出和增加公司利润的目的。毫无疑问，如果结果属实，那么这种做法并没有什么不对，毕竟禁令的初衷只是为了限制吸烟，而不是削减公司利润。


点拨

理性参与者从不吸烟。（玩笑话，这一点纯属本人虚构。）

3.1.2　弱占优

我认为，具备凡事无绝对的理念是理性的要素之一。

——伯特兰·罗素

让我们再一次回到意大利歌剧《托斯卡》的情境中。假设斯卡比亚在被托斯卡刺死之前并不在乎卡瓦拉多西

是死是活，那么相关的收益矩阵如图 3—6 所示。虽然托斯卡选择刺杀仍是占优策略（我们事实上也没有改变她的收益），但斯卡比亚选择下令发真子弹的策略却不再严格优于发空弹的策略（尽管该策略对他而言是一个理性的行动方案）。如果托斯卡屈服的话，斯卡比亚毫无疑问会下令发真子弹，这样做会使他的收益从之前的 3 增至 4。而假如托斯卡刺杀他，并且他本人也并不在乎，那么无论他如何行动，得到的收益只是 2。如此一来，下令发真子弹的策略不仅不会使他亏损，反而使他盈利。因而，下令发真子弹的策略弱优于发空弹的策略。

斯卡比亚

托斯卡	真子弹	空弹
刺杀	2,2	4,2
顺从	1,4	3,3

图 3—6　改编后的托斯卡博弈模型

定义 3.3

弱占优（weakly dominates）：若（1）无论其他参与者的策略是什么，s' 策略的收益至少等于 s'' 策略的收益；（2）对于其他参与者的一些策略，s' 策略的收益大于 s'' 策略的收益，则 s' 策略弱占优于 s'' 策略。

多数人的小心谨慎和对他人行为的预测缺乏绝对信心，使得不采用弱劣策略的做法显得很明智。这样做意味着你不会有所损失，反而可能会有所收益。不采用弱劣策略也不会有遗憾。如果存在弱占优策略（弱优于其他所有策略的策略），则使用该策略是一种明智的做法。



点拨

理性谨慎的参与者从不使用弱劣策略。理性谨慎的参与者总是采用弱占优策略。

3.1.3　拍卖会竞标

假设大学毕业 20 年之后，你以 5 000 万美元的价格卖掉了自己的网络公司。之后，你决定把这笔钱投向你所钟爱的现代艺术。此时，一家拍卖行将要拍卖一件你钟爱已久的拍品——艺术大师安迪 • 沃荷（Andy Warhol）的作品。参与竞标的条件是所有对该拍品感兴趣的人士必须在本周五下午 5 点之前提交一份书面竞标书。竞标价格最高者付款后将获得该拍品——这一方式被称为**首价拍卖**（first-price auction）。

假定你认为这件作品价值 40 万美元。如果你拍得这件作品，你的收益等于 40 万美元减去你支付的价格，而如果你没有竞拍成功，则收益为零。所以，你的支付价格等于 40 万美元时将没有收益；反之，如果你以高于（或低于）40 万美元的价格获得该拍品，则有收益（或亏损）。

你恰好得知自己只有一位竞争对手——自己大学时代的前女友。她在一家生物医药公司做了一段时间的首席执行官之后，最近刚在股市获利。你知道她对这件拍品的估价在 30 万美元，而且她也知道你的估价在 40 万美元。（在第 10 章当中，我们会研究每位竞拍者的估价不为对手所知的更为真实的情形。）

拍卖商告知拍品起价为 10 万美元，加价为每次 10 万美元。假设竞拍的最高价是 50 万美元。如果竞标价相同，则拍卖商投硬币决定获胜者。首价拍卖的策略组合如图 3—7 所示，图中策略和收益以 10 万美元计算。例如，如果你出价 3，她出价 1，那么你获胜，支付 3，并因此获得收益 1（=4−3）。如果你们都出价 3，那么你有 50% 的可能获胜，这种情况下支付 1 获得的收益是 3；你也有 50% 的可能失败，则导致收益为零，而预期收益为 $\frac{2}{3}$。（对于预期收益，我们会在第 7 章做进一步的阐释与说明。）

她

你	1	2	3	4	5
1	$\frac{3}{2}$,1	0,1	0,0	0,−1	0,−2
2	2,0	1,$\frac{1}{2}$	0,0	0,−1	0,−2
3	1,0	1,0	$\frac{1}{2}$,0	0,−1	0,−2
4	0,0	0,0	0,0	0,$-\frac{1}{2}$	0,−2
5	−1,0	−1,0	−1,0	−1,0	$-\frac{1}{2}$,−1

图 3—7　首价拍卖

你究竟应该出价多少呢？因为出价 5（超过了你的估价）严格劣于出价 4，显然，你不想出价那么多。而且，你也不想出价 4（你对拍品的估价），因为这一策略弱劣于其他更低拍价的策略。如果你以自己的估价出价，无论你是否得到拍品，你得到的收益都为零。那么，以较低的价格出价以期获得正收益似乎是更好的选择。这样一来，你可以剔除等于或高于自己估价的出价价格。此外，最低竞拍价 1 也属于弱劣策略。所以，我们不选择出价 1，4 和 5 的策略因为它们不是严格劣势策略就是弱劣策略。我们还可以进一步缩小范围吗？很可惜，答案是不能。出价 2 或出价 3 成为最佳策略的条件是由对手的出价决定。如果你认为她会出价 1，那么你就会出价 2；如果你猜测她会出价 3，你也会出同样的价格与之竞争。你想让竞拍价格低于自己的估价，但调低的幅度仍就取决于你对对手出价的预测。

你想调低出价的想法并不稀奇，因为在首价拍卖中，你只想让自己的出价略高于对手从而获得拍品（同时确保你的出价低于自己的估价）。成功的竞拍价格依赖于对手的出价。但是，我们惊奇地发现，拍卖方式略微调整之后，投标者的最佳竞拍价格竟与对手的出价毫无关联。经诺贝尔奖获得者威廉姆·维克瑞（William Vickrey）设计，改良的拍卖方式就如同首价拍卖一样——最高出价者获胜，但它又不同于首价拍卖，因为获胜者支付的不是自己的出价而是第二竞拍价格。这种拍卖方式被称为**二级价格拍卖**（second-price auction）。

二级价格拍卖的博弈策略组合如图 3—8 所示。例如，如果你出价 3，她出价 2，那么你获得拍品并支付 2（这是二级竞拍价格）。实际上，不论你出价 3，4 亦或是 5 得到的收益都是 2，因为你支付的价格不是自己的出价而是对手的出价。假如你打算出价 1，她计划出价 2，那么你的收益将因此受影响变为零，因为你较低的出价无法赢得拍品。

仔细观察图 3—8，我们会发现你出价 4 的策略弱占优于其他出价策略。所以，不论你猜测前女友会如何出价，选择出价 4 对你都是有利的。对她而言，出价 3 的策略弱占优于其他的出价选择。请注意你们每一个人的弱占优策略就是你们自己的估价。这一点并不是偶然的，因为在每一次的二级价格拍卖中，按估价出价的策略会弱占优于其他的出价策略。

你 \ 她	1	2	3	4	5
1	$\frac{3}{2}$,1	0,2	0,2	0,2	0,2
2	3,0	1,$\frac{1}{2}$	0,1	0,1	0,1
3	3,0	2,0	$\frac{1}{2}$,0	0,0	0,0
4	3,0	2,0	1,0	0,$-\frac{1}{2}$	0,−1
5	3,0	2,0	1,0	0,0	$-\frac{1}{2}$,−1

图 3—8　二级价格拍卖

在首价拍卖中，你使出价低于估价是为了在获得拍品的情形下尽可能降低需要支付的价格。但是这一策略在二级价格拍卖中并不起作用，因为你支付的不是自己的出价而是对手的竞拍价格。出价低于估价的策略只会降低你的成功率，所以并不是一个明智的选择。

计算出二级价格拍卖中的最佳出价是一件十分简单的事情。竞价者只需要判定拍品的价值，然后以此出价即可，既不需要一位执业心理医师帮助你评估其他竞拍者的心理，也不需要一位训练有素的博弈理论家告诉你如何竞拍。你需要做的只是了解自己。

案例研究

eBay 网上的委托出价悖论

你在 eBay 网上竞拍过物品吗？通常一次拍卖会持续数天，而 eBay 网的运作机制并不需要竞拍者每天在线 7 个小时，你甚至可以进入委托出价，其运作原理如下：只要其他竞价者的最高出价低于你的委托出价，你就是最高竞价者，竞拍价格等于其他竞拍者的最高出价加上最低加价。一旦其他竞价者的出价高于你的委托出价，你将失去最高出价（虽然你可以以更高的委托出价重新登上榜首）。

那么该如何确定你的委托出价？注意：如果在拍卖会接近尾声时，你递交了自己的最高委托出价，那么你将赢得拍品并且支付价格等于第二竞拍价格加上最低加价。如此一来，eBay 网上的拍卖倒有些类似于二级价格

拍卖。所以，你也应该相应地递交等同于自己估价的委托出价。此外，在递交了这样的出价之后，你可以在拍卖结束后回到拍卖现场看看自己是否赢得了拍品。很简单，是吧？

以自己的估价做委托出价的推论总是遭到现实的无情打击。人们在 eBay 网上拍卖的过程中经常会改变自己出价的做法，与理论上的预测大相径庭。举例来说，戴夫以 150 美元的委托出价竞拍一张滚石乐队的演唱会门票，一天之后，他发现最高出价已升至 180 美元（这样一来就超过了他的委托出价）。于是，他将自己的委托出价调至 200 美元。但是，如果起初戴夫就认为去看一场米克·贾格尔（Mick Jagger）和其同伴的演唱会值 200 美元，那他为什么不在一开始就提交 200 美元的委托出价？干嘛要在比估价低的委托出价上耽误工夫呢？

正是因为常常出现竞价者改变自己委托出价的现象，所以，我们不能草草地将这样的做法看成是“蠢人出价法”。这种现象代表了一种系统化的行为。而且，作为社会学研究者，我们的目的是要理解这种现象，而不是对此作出评判。如果我们认识到竞价者的行为方式受制于其思维方式，那么愚蠢的是理论本身；或者说得更好听一些，是我们的拍卖模式漏掉了一些相关要素。

究竟缺失了哪些要素呢？目前为止已经发现了一些可能的要素，但由于篇幅的关系这里只能列举其中之一。理论模式与现实之间的一个潜在偏差是 eBay 网对同一物品采用多重拍卖方式。仔细想想这一方式足以改变人们的出价策略。也许你认为一张滚石乐队演唱会的门票值 200 美元，这意味着你愿意以 200 美元的价格得到这样一张门票。但是，你在 eBay 网竞拍一张门票，竞拍失败并不意味着你无法得到这张票；相反，你可以参加滚石乐队演唱会门票的另一场拍卖。

为了理解这一新信息所产生的意义，假设你要参加一场于演唱会开始前两周结束的拍卖。假如你以 199 美元的拍价获胜，那么你的收益是 1 美元（等于你的估计减去出价）。虽然 1 美元的收益高于没有得到门票拍品的零收益，但不一定高于参加另一场拍卖的预期收益。为了在下场拍卖中以较低的出价赢得拍品，你可能会不愿意以 199 美元的拍价获胜。

不难理解这样的行动方案会使你随时改变自己的委托出价。假定你正在关注两场拍卖会：拍卖会 I 明天结束，拍卖会 II 两天后结束；而且你在拍卖会 II 递交了委托出价，对于拍卖会 I 的出价持观望态度。正如刚才的论证，你的委托出价并不是你的估价，而是基于你认为能够让自己在另外一场拍卖会获胜的适当价格。如果拍卖会 I 的最高出价高于你的预期，那么你就会认为剩下门票的出价会更高，从而提高你在拍卖会 II 中的委托出价。你的最佳委托出价会随着这些门票在 eBay 网其他拍卖中价格的攀升而改变。

eBay 网上拍卖中理论预测与实际行为脱节的现象揭示出某一特定博弈理论模型的不足，而不是整个博弈理论。后者不仅具有极强的可塑性，并且可以和富有观察力的智者一起对许多社会现象作出令人信服的解释。

3.1 理解测验

根据图 3–9 所示的博弈情形，求出哪些是严格劣势策略，哪些是弱劣势策略。

		参与者2			
		w	x	y	z
参与者1	*a*	3,2	1,1	4,3	3,5
	b	1,3	3,0	2,4	4,2
	c	2,1	0,1	1,2	1,0
	d	1,0	2,0	2,1	4,0

图 3—9

3.2 博弈求解：参与者是理性的和参与者知道参与者是理性的

> （萨达姆·侯赛因）一开始就表现一副咄咄逼人的样子，这让人有些退缩。我记得当时看到自己的双手在冒虚汗。我很清楚他知道自己有着怎样的名声，而他也知道我了解他的名声如何。
>
> ——比尔·理查森，美国政治家

在二级价格拍卖中，参与者可以在无需了解其他参与者出价策略的前提下选择最佳出价策略。然而，在首价拍卖中，出价低于自己估价的幅度取决于你对其他竞价者最高出价的预测。博弈通常要求具有类似于首价拍卖的竞拍思维方式，即参与者必须预测其他参与者的行为。

本节中，通过对几个案例的研究，我们发现参与

者是理性的设想并不足以求出博弈的解。但是，如果进一步假设——比如每个参与者都知道其他参与者是理性的，那么我们就可以对参与者的行为表现（至少是在某些情境中的行为表现）得出合理的结论。

案例研究

团队合作

> 斯坦福高中无论是学生数量还是人才类型方面都是数一数二的。这其中既有运动员，也有书呆子；不仅有女子社团，也包括男子联谊会；此外还有独立青年、预科学生等，真可谓穷尽了各类学生。
>
> ——《斯坦福日报》(*Stanford Daily*)

假设大学某个班级的学生形形色色，各不相同，而我们只挑选其中的几种类型。类型一：学习成绩不良的运动员。因为很少花时间学习，所以他们很满足于得到C等成绩（在我的母校弗吉尼亚大学，这类成绩因其看上去像一个挂钩而被称为“hook”）。类型二：满足于B等成绩的社团成员，但他们更愿意通过努力学习避免成绩下滑和令父母不满。类型三：学习成绩优异的书呆子。他们努力学习就为了获得A等成绩，对于他们而言，没有什么比埋头苦读更好（还有没遭到批判的学生类型吗）。

根据学生任务（比如考试）的完成情况判断其学习努力程度并不是一件难事。书呆子刻苦学习，社团成员适度学习，运动员的学习程度刚好能够通过考试。当把他们放在一起组合成一个团体会是一种怎样的情形呢？团队的质量和成绩由团队所有成员共同决定。一个学生的努力程度很可能取决于团队其他成员期待努力的程度。

为了便于分析，假定团队由两人组成。先研究由书呆子和运动员组成的团队，其联合收益矩阵如图3—10所示。每一个学生都有三种努力程度：较少努力、适度努力和非常努力。假定团队成绩会随着两人的努力而提高，那么一个学生的收益总是会因为另一个学生的努力而有所增加，因为另一个学生的努力（不需要更加努力）而提高成绩。

		书呆子		
		较少努力	适度努力	非常努力
运动员	较少努力	3,1	4,2	5,3
	适度努力	2,2	3,3	4,4
	非常努力	1,3	2,4	3,5

图3—10　书呆子和运动员组成的团队合作博弈

运动员非常不喜欢学习，对他们的收益进行高低排序可以反映出不愿努力学习的倾向。无论她（没错，当然包括女运动员）的同伴——书呆子付出多大的努力，运动员的收益会因为书呆子的越发努力而有所减少。比如，如果书呆子适度努力，那么运动员的收益就会随着其同伴由低、中到高的努力程度从4降至3，再降至2。所以，采用最低努力程度的策略成为了运动员的占优策略，因为不管她的同伴如何努力，较少努力的策略产生的收益要高于其他的策略。

书呆子的情况又如何呢？他的收益会随着努力的增加而增加。不管他的同伴是否努力，他都会选择更加努力从而提高团队的成绩。因而，书呆子的占优策略是选择非常努力。图3—10显示的博弈结果清楚地表明：如果学生是理性的，那么运动员会选择较少努力的策略，书呆子会采用非常努力的策略。正是因为和愿意努力的同伴在一起，所以运动员能够得到较高的收益5，而她的同伴书呆子得到的收益是3（尽管书呆子会喃喃抱怨其同伴是个“愚蠢懒散的家伙”）。

接下来，让我们关注一下由社团男生和书呆子组成的团队（其收益矩阵见图3—11）。如之前所述，书呆子的收益会随着努力的增加而增加。而社团男生的情况要比书呆子和运动员的复杂一些。他希望获得一个中等成绩并且愿意努力去达到这样的要求，但却不会继续努力将成绩由B等提升至A等。这样，他就没有一个占优策略。如果他的同伴选择较少努力，那么他会努力去获得B等成绩。如果他的同伴“全力以赴”，那么他会满足于维持现状（因为此时他仍能获得B等成绩）。而如果他的同伴只是适度努力，那么他也会这样做。

由此看出，仅仅假设社团男生是理性的并不能帮助我们预测他的行为选择。如果我们假定不仅参与者是理性的，能求出这一博弈的解吗？别忘记这一博弈的共同知识是社团男生知道自己和书呆子是一个团队（这一点从对方眼镜架上的胶带就可以看出来）。设想一下社团

男生不仅是理性的，而且知道他的同伴也是理性的。通常理性人士会在具备占优策略的条件下使用该策略，因而社团男生推测因为他的同伴是理性的，所以他会非常努力学习。鉴于此，社团男生会选择较少努力的策略。因此，在由社团男生和书呆子组成的团队中，书呆子会忙得不可开交，而社团男生会闲得发慌。为了得出这一结论，我们需要假设书呆子和社团男生是理性的，同时社团男生知道他的同伴书呆子是理性的。

书呆子

		较少努力	适度努力	非常努力
社团男生	较少努力	0,1	2,2	6,3
	适度努力	1,2	4,3	5,4
	非常努力	2,3	3,4	3,5

图 3—11　书呆子和社团男生组成的团队合作博弈

最后，把我们的注意力转向社团男生和社团女生的组合（其收益矩阵如图 3—12 所示）。假定参与者是理性的和每一个参与者知道其他参与者是理性的并不足以求出该博弈的解。适用于社团男生和书呆子组合博弈的方法之所以在这里行不通，是因为社团男生和女生都没有一个占优策略。欲求解该博弈请听第 4 章分解。

社团女生

		较少努力	适度努力	非常努力
社团男生	较少努力	0,0	2,1	6,2
	适度努力	1,2	4,4	5,3
	非常努力	2,6	3,5	3,3

图 3—12　社团男生和社团女生组成的团队合作博弈

上帝存在论

> 人类无法证明上帝的存在正如同哈利·波特无法证明小说家乔安娜·凯瑟琳·罗琳的存在一样。
>
> ——弗雷德里克·布赫纳

哲学家们对于上帝是否存在一直争论不休。其中最著名的一个观点是由布莱斯·帕斯卡（Blaise Pascal）在 17 世纪中叶提出的。布莱斯·帕斯卡是一位天才数学家，他曾经有一段时间放弃了数学转而投身于上帝。他对上帝存在论的解释被人们称为帕斯卡赌注。内容如下：

> 假设你不确信上帝的存在。倘若你不相信上帝的存在，而事实上上帝确实存在，那么你将受到严厉的处罚（想象一下炽热的干草叉和被一轮又一轮捆扎的情景）。如果你相信上帝的存在，但事实是上帝并不存在，那么你付出的代价将很小。

因此，帕斯卡主张人们最好相信上帝的存在，否则等待他的将会是极其可怕的酷刑。

大家对于帕斯卡赌注的评价不一。人们真的能够选择自己的信仰吗？上帝奖赏的是信仰还是行动？信仰上帝应该出于冷静的推理选择而不是忠诚或爱吗？由于本书关注的是博弈理论而非哲学或神学，所以我认为帕斯卡赌注的问题在于只有一位决策者。是否可以将上帝也看做博弈当中的一位参与者呢？特别是，如果上帝有权决定是否在人类面前显示自己的存在的话，上帝会怎么做，在这种情形之下人类又会怎么做？

此时，上帝有两种策略：显示或隐藏自己。人类有帕斯卡提出的两种策略：相信或不相信上帝的存在。为了方便描述收益状况，假设人类最关心的是证明自己相信（或不相信）上帝存在的事实。也就是说，倘若他相信上帝的存在，则他想看到上帝存在的事实；如果他不相信上帝的存在，那么他当然不想看到上帝存在的事实。其次才是人们愿意相信上帝的存在。对上帝而言，她最关注的是人们对上帝的信仰，然后才是隐藏自己。该博弈的策略组合如图 3—13 所示。

人类没有占优策略。如果上帝打算现身，则人类愿意相信上帝的存在。倘若上帝想隐身，那么人类不愿相信上帝的存在。相反，上帝有占优策略：不管人类是否相信上帝的存在，她都选择隐藏自己。所以人类选择相信上帝存在时，上帝不隐藏与隐藏的收益比是 4:3；人类选择不相信上帝存在时，上帝不隐藏与隐藏的收益比是 1:2。理性的上帝会选择隐藏自己。

人类

		相信	不相信
上帝	现身	3,4	1,1
	隐身	4,2	2,3

图 3—13　上帝存在论博弈

假如人类相信上帝是理性的，那么人类知道上帝会

隐藏自己（因为这是她的占优策略）。既然上帝会选择隐藏自己的策略，则人类的最佳策略就是不相信上帝的存在。我们认为这一谜团的答案是人类不应该相信上帝的存在，上帝也从不现身。①

智猪博弈

本书旨在揭示如何借助博弈理论理解人类行为，但该理论同样可以用来解释低等动物的行为（详见第16章和第17章）。假设猪圈里有一头大猪，一头小猪。猪圈的一头安装有一个杠杆，另一头装有一个食物槽。只要一踩杠杆，另一头的食物槽里就将流出十份食物。踩杠杆需要花费的能量，相当于消耗两份食物。十份食物的分配取决于哪头猪先到食物槽以及它的体形大小。如果大猪先到，则它可以得到九份食物，小猪只得到一份。大猪不仅具有体重优势，而且在位置上也占优（想象一下篮球场上篮球运动员沙奎尔·奥尼尔与歌王迈克尔·杰克逊对峙得分的情景）。相反，倘若小猪先到，则得到四份食物（因为在被大猪赶走之前已经吃到了一些）。如果它们同时到达，预计小猪能得到三分食物（很可能是大猪嘴里漏掉的那些食物）。

每头猪决定是否踩杠杆的选择就是它们的两种备选策略。假设一头猪的收益是得到的食物减去踩杠杆消耗的食物，则图3—14显示了该博弈的策略组合。

		大猪	
		踩杠杆	在食物槽边等待
小猪	踩杠杆	1,5	−1,9
	在食物槽边等待	4,4	0,0

图3—14 智猪博弈

答案是大猪不踩杠杆而坐享其成吗？并非如此。在这种情况下，“弱势就是力量”。因而，答案是大猪去踩杠杆而小猪则在另一头坐享其成。为什么会出现这样的结果？

问题的关键就在于小猪有占优策略。如果大猪踩杠杆，小猪最好在食物槽边等待，这样做不仅可以得到更多的食物而且可以减少踩杠杆所消耗的食物，因此得到收益4，远远大于踩杠杆得到的收益1。倘若大猪在食物槽边等待，则导致小猪得到的食物不足以满足踩杠杆所消耗的食物，所以得到的收益是−1（即得到的一份食物减去踩杠杆消耗的两份食物）。相比之下，选择不去踩杠杆的策略（收益为零）显得更为明智。因此，若小猪是理性的（尽管听起来可能有点怪），则不论大猪怎么做，它都会选择在食物槽边等待。

由于大猪没有占优策略，所以如果小猪踩杠杆，它会在食物槽边等待；如果小猪在食物槽边等待，它就会去踩杠杆。假设大猪相信小猪是理性的（当然听起来很奇怪），那么它知道小猪会在食物槽边等待。意识到小猪踩杠杆得到的食物满足不了踩杠杆所消耗的食物，大猪相信小猪不会去踩杠杆。在这种情境下，大猪只好去踩杠杆（尽管这样做意味着小猪会有先得到食物的优势）。无疑大猪一到达食物槽边，就会放开肚皮大吃一顿以补充踩杠杆所消耗的食物。所以如果这两头猪是理性的，并都相信对方是理性的，那么博弈的解就是大猪踩杠杆，小猪在食物槽边等待。

假设猪是理性的就等同于猪会飞吗？事实上，假设猪是理性的并非不合理，因为这意味着猪会以自己的利益行事。从这一点来说，所有的物种不都是如此吗？更具争议性的是假定猪认为其他的猪是理性的，这看上去确实有些不靠谱。但是在你放弃这个解决方案之前，我们不妨先按实际情境中猪的行为表现得出另一解决方案，两者对比之后，再来看该方案是否可行。

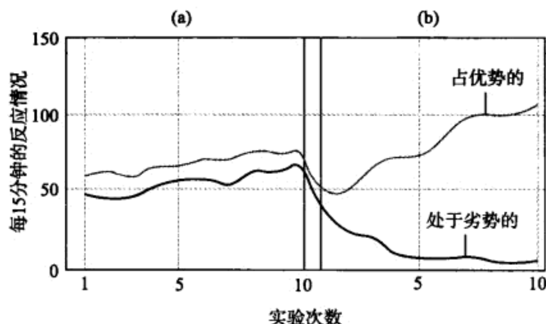
实验笼子的尺寸是2.8米×1.8米。为了确保这两头猪具有很强的食欲，连续24小时不让其进食。起初每头猪被单独放在笼子里，以训练其意识到踩杠杆会得到食物。实验者通过将两头猪放在一个有一碗食物的猪圈里来评判优劣，不再将体型大小作为决定性因素。如果一头猪大部分时间只是从碗里取食的话，就被认为是占优的。

实验结果如图3—15所示。图中的竖坐标表示每十五分钟杠杆被踩动的次数，横坐标表示的是尝试的次数。尝试踩杠杆的次数达到10之前，这两头猪被单独放在笼子里，占优的猪踩杠杆的次数略多一些。10次尝试之后，它们被放在同一个笼子里，结果很令人吃惊：占优的猪踩杠杆的次数越来越多。

我认为这两头猪取得这样的结果并不是因为一头猪

①不要将话题转向理性（人类特有的概念）是否适用于上帝，人类如何与自己都不确信是否存在的人进行博弈，得出人类不应该也不会相信上帝存在的结论是否亵渎上帝等方面。我们举这个例子只是为了启发人们如何创新应用博弈理论。——作者注

能够猜测另一头猪的想法。相比之下，它们通过不断的尝试变得聪明起来也许是更恰当的解释。注意观察，它们的行为也确实越来越接近于预测的结果。也许处于劣势的猪在几次踩踏杠杆却只得到很少食物之后，不再踩杠杆。这时，占优势的猪会发现它能吃到食物的唯一途径就是踩杠杆。经验有时可以很好地替代机智推理。



图中记录的是一头占优势的猪和一头处于劣势的猪在 24 小时未进食的情况下每 15 分钟踩杠杆的平均次数，其中（a）是两头猪分开的实验结果，（b）是两头猪在一起的实验结果。

图 3—15　智猪博弈实验结果。

3.2 理解测验

按照图 3–16 所示的博弈情形，解出与参与者是理性的和每个参与者相信其他参与者是理性的相一致的策略。

参与者2

参与者1		*w*	*x*	*y*	*z*
	a	3,2	1,1	4,3	3,5
	b	1,3	3,0	2,4	1,5
	c	2,1	0,1	1,2	1,0
	d	1,0	2,0	2,1	4,2

图 3—16

3.3 博弈求解：理性成为共同知识

黑衣人：好吧，毒药在哪？这场游戏才刚刚开始。你先挑，然后我们同时喝下，不管谁生，谁死，游戏都将结束。

威兹尼：这太简单了吧。我只要猜测你的想法就行。你喜欢把毒药放在自己的杯子里还是对手的杯子里？聪明的人总是会把毒药放在自己的杯子里，因为他知道只有大傻瓜才会选择自己眼前的东西。我不是傻瓜，所以我不会选择你面前的酒。但是，你一定知道我不是大傻瓜，那么我当然也不会选择放在自己面前的酒。

——节选自电影《公主新娘》

（*The Princess Bride*）

3.3.1 兴奋剂博弈：运动员服用类固醇是否理性

2007 年 8 月 7 日，棒球运动员贝瑞•邦兹（Barry Bonds）打出第 756 支本垒打，打破了由亨利•阿伦（Henry Aaron）创造的 755 支本垒打纪录。对于邦兹而言，这本是一个值得骄傲和受人赞美的时刻，但这一成就却因为不断受到指控而蒙上阴影。这些指控称他的优异表现并不是努力的结果而是与服用类固醇有关。最近出版的一本名为《影子游戏》（*Game of Shadows*）的书介绍说，邦兹从 1998 年就开始服用类固醇。

不论邦兹是否服用类固醇，可以肯定的一点是兴奋剂问题不仅存在于专业运动领域，就连业余体育也深受影响。为了应对这一挑战，我们应尽快制定一个体系以帮助运动员远离兴奋剂，这不仅是对运动爱好者负责，更是让运动员自身受益。服用类固醇是为了让某一运动员略优于其他运动员，但是如果所有的运动员都服用类固醇，那么优势也就不存在了。利益消失了，但运动员仍需付出损害身体健康的代价。博弈理论可以用来研究如何构建一个监督惩戒机制，从而提供积极正确的引导。

虽然本节并不着手进行这一极富挑战性的任务，但是至少可以分析出运动员所面临的各种诱惑以及这些诱惑如何影响他们的行为方式。设想有三名颇具竞争力的世界级短跑运动员。他们的天赋和服用类固醇的倾向（这一点由获胜的欲望判定）各不相同。在天赋方面，假定莫里斯比卡尔快，卡尔比本快。在服用类固醇的倾向方面，本的倾向性高于卡尔，卡尔高于莫里斯。具体而言，无论卡尔和莫里斯是否服用类固醇，本都会服用。如果没人使用类固醇，卡尔也不会，但是为了富有竞争力，所以假如本或者莫里斯亦或是两个人都使用类固醇，那么卡尔就会服用。在不服用类固醇的前提下，莫里斯是三者当中跑得最快的。然而，如果本和卡尔都选择服用

类固醇，那么莫里斯也会服用。

图 3—17 表示的是服用类固醇倾向性的策略式博弈。表格中的第一个数字是莫里斯的收益，第二个数字是卡尔的收益。莫里斯位于横栏，卡尔位于竖列，本选择矩阵。

这些运动员会作何选择？理性无法预测卡尔和莫里斯的选择，因为他们服用或放弃类固醇的策略取决于对其他运动员策略的预测。不过，服用类固醇是本的占优策略（见表 3—18）。倘若卡尔和莫里斯都不使用类固醇，则本使用类固醇的收益是 6，高于不使用类固醇的收益 4。若或者卡尔或者莫里斯使用类固醇，则本服用类固醇的收益是 5，不使用的收益是 2。最后，假使卡尔和莫里斯都使用类固醇，那么本使用类固醇的收益是 3，放弃使用的收益是 1。因而，理性意味着本会使用类固醇。

本服用类固醇

莫里斯 \ 卡尔	服用类固醇	不服用类固醇
服用类固醇	2,3,3	3,1,5
不服用类固醇	1,4,5	5,2,6

本不服用类固醇

莫里斯 \ 卡尔	服用类固醇	不服用类固醇
服用类固醇	3,4,1	4,2,2
不服用类固醇	5,5,2	6,6,4

图 3—17　兴奋剂博弈

本服用类固醇

莫里斯 \ 卡尔	服用类固醇	不服用类固醇
服用类固醇	2,3,③	3,1,⑤
不服用类固醇	1,4,⑤	5,2,⑥

本不服用类固醇

莫里斯 \ 卡尔	服用类固醇	不服用类固醇
服用类固醇	3,4,①	4,2,②
不服用类固醇	5,5,②	6,6,④

图 3—18　服用类固醇是本的占优策略

让我们假设不仅这些运动员是理性的，而且每个人相信其他两名运动员是理性的。这一假设意味着卡尔和莫里斯都相信本会使用类固醇，因为本的理性是使用类固醇。对于卡尔和莫里斯而言，博弈的策略形式如图 3—19 所示，当中去掉了本放弃类固醇的策略。卡尔拥有一个服用类固醇的占优策略。既然卡尔知道本会使用类固醇（因为他知道本是理性的并且服用类固醇是本的占优策略），那么卡尔也会使用类固醇（因为无论莫里斯是否服用类固醇，这都是他的最佳策略）。莫里斯仍然没有占优策略。

那么到目前为止，已知本会使用类固醇（因为他是理性的），卡尔会使用类固醇（因为卡尔是理性的并且他知道本是理性的）。在这种情况下，莫里斯会怎么做？设想每个运动员知道所有运动员都知道大家是理性的。这一假设告诉我们莫里斯知道卡尔了解本是理性的。因而，莫里斯不仅知道卡尔了解本会使用类固醇，而且知道卡尔会使用类固醇。那么，莫里斯排除卡尔不使用类固醇的策略，他所处的形式如图 3—20 所示。鉴于莫里斯推测本和卡尔都使用类固醇，他认为使用类固醇（产生的收益是 2，高于放弃使用类固醇策略的收益 1）也是自己的最佳策略。

莫里斯 \ 卡尔	服用类固醇	不服用类固醇
服用类固醇	2,③,3	3,①,5
不服用类固醇	1,④,5	5,②,6

此时，莫里斯可以推测出卡尔的占优策略是服用类固醇

图 3—19　本服用类固醇时的博弈

莫里斯	服用类固醇
服用类固醇	2,3,3
不服用类固醇	1,4,5

图 3—20　本和卡尔都服用类固醇的博弈情形

我们得出如下结论：倘若（1）所有运动员是理性的，（2）每个运动员都相信其他运动员是理性的，（3）每个运动员都相信其他运动员相信其他运动员是理性的，则这三名运动员都使用类固醇。这一结论中令人沮丧的一点是当中有两名运动员并不想使用类固醇，只是因为其他运动员使用类固醇而不得不这样做。本想借助于化学药品提升自身成绩的强烈倾向导致其他两名运动员也屈服于这样的诱惑。这正是现今的体育界面临的挑战。

该结论听上去真实可信。《影子游戏》一书的作者声称邦兹是在 1998 年赛季之后才开始服用类固醇的。当时，备受世人关注的马克·麦奎尔（Mork McGwire）和山米·苏萨（Sammy Sosa）之间竞争激烈，两人都力图打破罗杰·马里斯（Roger Maris）创造的单赛季 61 支本垒打的记录。事实上，两人也确实做到了。被指称服用类固醇的马克·麦奎尔创造出单赛季 70 支本垒打的新纪录。三年后，邦兹以 73 支本垒打打破了他的纪录。如果邦兹确实使用类固醇，不也是为了在与其他顶级棒球击球手的激烈竞争中保持竞争力吗？

谜题

假设每个运动员都是理性的并且每个人相信其他运动员相信他是理性的。但是，每个运动员不确信其他运动员是否理性。比如，卡尔是理性的并且相信莫里斯和本都认为他是理性的，但是卡尔不知道莫里斯和本是否是理性的。那么这些信念有意义吗？你认为运动员会怎么做？

3.3.2 重复剔除严格劣势策略

到目前为止，我们已经通过逐步增加理性知识来求得博弈的解。《托斯卡》博弈只假设参与者是理性的就得出了解。上帝存在论博弈的假设不仅包括参与者是理性的，也包括每个参与者相信所有的参与者都是理性的。具体来说就是需要假设上帝和人都是理性的，人相信上帝是理性的。兴奋剂博弈中，运动员是否使用类固醇策略的解是假设所有的参与者都被认为是理性的，每个参与者相信所有的参与者都是理性的，以及每个参与者相信所有参与者相信所有参与者是理性的后得出的。

这些都是博弈求解的一个更为普遍的准则的例证，该准则被称为**重复剔除严格劣势策略**（IDSDS）。

重复剔除严格劣势策略准则可以经由以下一连串步骤定义。

- 步骤 1：剔除原有博弈中的所有严格劣势策略（这一步是基于所有参与者都是理性的假设）。
- 步骤 2：剔除完成第一步后博弈中的所有严格劣势策略（这一步是基于每个参与者相信所有参与者都是理性的假设）。
- 步骤 3：剔除完成第二步后博弈中的所有严格劣势策略（这一步是基于每个参与者相信所有参与者相信所有参与者都是理性的假设）。
- 步骤 4：剔除完成第三步后博弈中的所有严格劣势策略（这一步是基于每个参与者相信所有参与者相信所有参与者相信所有参与者都是理性的假设）。

⋮

- 步骤 t：剔除完成 t-1 步后博弈中的所有严格劣势策略。

这个过程直到没有策略可以剔除才会结束。如果某一博弈中每个参与者有无限个策略，则这个过程就会无限进行下去，但这种情况很少见。通常，在经过有限的步骤之后，就没有可以剔除的策略了，剩下的就是可以在重复剔除严格劣势策略准则中继续使用的策略。

以本章一开始引用的夏洛克·福尔摩斯的话为例，我们发现重复剔除严格劣势策略准则剔除了“不可能”，那么不论剩下什么都是可能的。如果每个参与者只剩一个策略（注意：至少得剩下一个策略），则博弈是**占优可解的**（do-minance solvable），而且重复剔除严格劣势策略准则对行为的预测具有唯一性。

让我们通过例证分析来加深对这个过程的理解。以图 3—21 所示的两人博弈为例。

参与者2

参与者1	w	x	y	z
a	3,2	4,1	2,3	0,4
b	4,4	2,5	1,2	0,4
c	1,3	3,1	3,1	4,2
d	5,1	3,1	2,3	1,4

图 3—21　重复剔除严格劣势策略的运用

先关注一下步骤 1 中的参与者 1，她有严格劣势策略吗？要回答这个问题，你可以考虑每个策略，然后判断对于参与者 2 的每个策略，是否有另一个策略可以产生一个严格劣势策略。一个简便的办法是先判断对于参与者 2 的一些策略，哪些策略是参与者 1 的最佳策略。这些策略不是严格劣势策略，因为他们在某种情形下是最优的。

先确定参与者 1 的最佳策略技巧。注意假如参与者 2 选择策略 *w*，则策略 *d* 是参与者 1 的最佳策略（得到的收益是 5，高于其他任何策略的收益）。这样，*d* 就不是严格劣势策略。若参与者 2 使用策略 *x*，则策略 *a* 是参与者 1 的最佳策略，因而策略 *a* 不是严格劣势策略。参与者 2 使用策略 *y*，那么策略 *c* 就是参与者 *a* 的最佳策略，所以也不是严格劣势策略。最后，倘若参与者 2 使用策略 *z*，则策略 *c* 是参与者 *a* 的最佳策略，但已知 *c* 不是严格劣势策略。到目前为止，可知策略 *a*、*c* 和 *d* 不是参与者 1 的严格劣势策略，那么可以考虑的只有策略 *b*。事实上，策略 *b* 严格劣于策略 *d*。所以，正如图 3—22 所描述的，我们可以剔除图 3—21 中的策略 *b*。

参与者2

参与者1	w	x	y	z
a	3,2	4,1	2,3	0,4
~~b~~	~~4,4~~	~~2,5~~	~~1,2~~	~~0,4~~
c	1,3	3,1	3,1	4,2
d	5,1	3,1	2,3	1,4

图 3—22　剔除参与者 1 的严格劣势策略后的博弈情形

步骤 1 并没有结束，因为还需找出参与者 2 的最佳策略。仍从图 3—21 中的博弈策略组合开始，如果参与者 1 选择策略 *a*，则策略 *z* 是参与者 2 的最佳策略，所以 *z* 不是严格劣势策略。若参与者 1 选择策略 *b*，则策略 *x* 是参与者 2 的最佳策略，那么 *x* 不是严格劣势策略。倘若参与者 1 选择策略 *c*，则策略 *w* 是参与者 2 的最佳策略，这样，*w* 不是严格劣势策略。如果参与者 1 选择策略 *d*，那么 *z* 又是参与者 2 的最佳策略。因而，策略 *w*、*x* 和 *z* 都不是严格劣势策略。剩下的策略 *y* 严格劣于策略 *z*，因为参与者 2 是理性的，所以我们认为他不会使用策略 *y*。这样一来就可以剔除掉策略 *y*（参见图 3—23）。

把目光转向步骤 2，图 3—24 表示的是精炼后的博弈，其中剔除了参与者 1 的策略 *b* 和参与者 2 的策略 *y*。还有可以剔除的严格劣势策略吗？参与者 1 没有可以剔除的严格劣势策略（自己想想为什么）。对于参与者 2，策略 *z* 严格优于策略 *x*。注意原有博弈中策略 *x* 并不严格劣于策略 *z*，因为当参与者 1 使用策略 *b* 时，策略 *x* 产生的收益要高于策略 *z*（和其他任何一个策略）的收益。不过，因为策略 *b* 是参与者 1 的严格劣势策略，参与者 2 不相信参与者 1 会使用该策略（参与者 2 认为参与者 1 是理性的），所以，已经剔除策略 *b*，同时将策略 *x* 看成是一个潜在的有益策略。参与者 2 的其他策略仍是非劣策略。

参与者2

参与者1	w	x	~~y~~	z
a	3,2	4,1	~~2,3~~	0,4
~~b~~	~~4,4~~	~~2,5~~	~~1,2~~	~~0,4~~
c	1,3	3,1	~~3,1~~	4,2
d	5,1	3,1	~~2,3~~	1,4

图 3—23　剔除参与者 2 的严格劣势策略后的博弈

参与者2

参与者1	w	x	z
a	3,2	4,1	0,4
c	1,3	3,1	4,2
d	5,1	3,1	1,4

图 3—24　经过一轮重复剔除严格劣势策略后的博弈情形

虽然步骤 2 中剔除了一个策略，但该过程并未结束，需要进入步骤 3。图 3—25 所示的是剔除参与者 2 的策略 *x* 后的博弈情况（记住没有剔除参与者 1 的策略。）。细心观察之后，我们发现对于参与者 1 而言，策略 *d* 严格优于策略 *a*，策略 *c* 和 *d* 是非劣策略。显然，策略 *w* 和 *z* 仍然是非严格劣势策略：在步骤 2 中，它们不是严格劣势策略，参与者 1 在步骤 3 中的策略与步骤 2 中的相同。

	参与者2	
参与者1	*w*	*z*
a	3,2	0,4
c	1,3	4,2
d	5,1	1,4

图 3—25　经过两轮重复剔除严格劣势策略后的博弈情形

剔除策略 *a* 后的博弈情形如图 3—26 所示，由此进入步骤 4（此时没有一个策略是严格劣势策略）。由于没有可剔除的策略，所以该过程结束。因而，得出的结论是经过重复剔除之后剩下的策略是参与者 1 的策略 *c*、*d* 和参与者 2 的策略 *w*、*z*。那么，理性是共同知识的假设虽然不能做出唯一的推测，但至少可以帮助我们从 16 对备选的策略中剔除掉 12 对，即参与者 1 会使用策略 *c* 或 *d*，参与者 2 会使用策略 *w* 或 *z*。

	参与者2	
参与者1	*w*	*z*
c	1,3	4,2
d	5,1	1,4

图 3—26　经过三轮重复剔除严格劣势策略后的博弈情形

3.3 理解测验

按照图 3.16 所示的博弈情形，求出重复剔除严格劣势策略后所剩的策略。

本章总结与练习题

本章总结

本章概述了如何在参与者是理性的，即参与者 1 知道参与者 2 是理性的，参与者 3 知道参与者 1 知道参与者 2 是理性的等条件下，求出博弈的解。需要强调的是理性参与者不使用**严格劣势策略**。不管其他参与者如何选择，如果现有的策略产生的收益小于另一策略，那么该策略就是严格劣势策略。所以，无论参与者对其他参与者的策略有什么样的信念，他都不会选择严格劣势策略。正是由于**占优策略**严格优于其他任何一个策略，所以它无疑是博弈中最具竞争力的策略。

不使用**弱劣策略**也是理性的表现之一。倘若另一策略和现有策略的收益相同，甚至有时优于现有策略（譬如对于其他参与者的一些策略而言），那么该策略就是弱劣策略。使用弱劣策略只会令你后悔。在某些博弈（如二级价格拍卖博弈）中，使用**弱占优策略**是理性参与者的不错选择。

重复剔除严格劣势策略（IDSDS）准则是基于理性参与者不使用严格劣势策略的理念产生的。该准则的核心是参与者是理性的以及参与者不使用严格劣势策略是共同知识的假设。重复剔除严格劣势策略准则剔除博弈中每个参与者的严格劣势策略，以减少博弈中的备选策略（只有严格劣势策略才可以被剔除）。在精炼后的博弈中，严格劣势策略再一次被剔除。某一策略在原有博弈中可能并不是严格劣势策略，但在剔除其他参与者的一些策略后会成为严格劣势策略。剔除严格劣势策略的过程会持续到剩下的策略都不是严格劣势策略为止。经过重复剔除之后剩下的策略就是可行的解决方案。

就某些博弈（如《托斯卡》博弈）而言，参与者是理性的假设可以推导出一个具有唯一性的**策略组合**（strategy profile），但前提是每个参与者都有一个占优策略。而对于包括上帝存在论博弈在内的其他博弈而言，要想求出唯一的解不仅需要假设所有参与者是理性的，还需假设所有参与者知道所有参与者是理性的。当然，重复剔除**糊涂策略**（stupid strategy，即严格劣势策略）的过程只有在原有博弈中存在糊涂策略时才可行。事实上，许多博弈情形中都没有糊涂策略（我们会在下一章讨论该策略），因而，重复剔除严格劣势策略的准则是无法剔除备选策略组合的。

本章介绍了一些由博弈理论提出的颇具洞察力的见解。譬如，《托斯卡》博弈中，每个参与者的个人

理性是如何影响集体利益的；香烟广告博弈中，参与者的选择越少收益越高；智猪博弈中，小猪如何战胜大猪。在以后的章节中，我们会介绍更多独到的见解。

细心的读者会发现我们并没有求出本章开始时图3—1所示博弈的解。尽管运用本章介绍的方法无法预测绑架博弈的结果（因为不存在严格劣势策略），但幸运的是，博弈理论总是有许多的解决方法，其中的另一个方法会在下一章中予以说明。拿娱乐界的行话来说就是，总要吊着人们的胃口。

附录1和附录2进一步阐释了严格占优和劣占优概念，以及先进理念——可理性化。

练习题

1. 求出第2章中图2—8所示的抢劫博弈的策略式，然后推断是否有一些策略是严格劣势策略或弱劣策略？

2. 在苏斯博士讲述的“扎克斯”故事中，有一个向北走的“扎克斯”和一个向南走的“扎克斯”，两人在行进过程中碰面了。为了让对方顺利通过，每个人必须决定是按原有的方向走下去还是变换方向。结果，没有人移动且这样的困境持续了很长时间。请写出这一情境中博弈的策略式。

3. 在团队合作博弈中，假设运动员和社团女生组成一个团队（如图3—27所示）。

		社团女生		
		较少努力	适度努力	非常努力
社团男生	较少努力	3,0	4,1	5,2
	适度努力	2,2	3,4	4,3
	非常努力	1,6	2,5	3,4

图3—27

（1）如果两个人都是理性的并且运动员知道社团女生是理性的，那么结果会如何？

（2）倘若两个人都是理性的并且社团女生知道运动员是理性的，那么结果是什么？

4. 仔细观察图3—28所示的策略式博弈。

a. 假如两位参与者都是理性的，结果如何？

b. 假如两位参与者是理性的并且每位参与者相信对方是理性的，会出现什么结果？

c. 找出重复剔除严格劣势策略后所剩的策略。

		参与者2		
		x	*y*	*z*
参与者1	*a*	1,3	1,1	0,2
	b	3,1	2,2	1,0
	c	0,2	1,2	3,0

图3—28

5. 按照图3—29所示的策略式博弈，推出重复剔除严格劣势策略后所剩的策略。

		参与者2		
		x	*y*	*z*
参与者1	*a*	5,2	3,4	2,1
	b	4,4	3,2	3,3
	c	3,5	4,4	0,4
	d	2,3	1,5	3,0

图3—29

6. 假设两个盖尔特部落——“Garbh”部落和“Conchubhair”部落（这两个部落的名字你想怎么念就怎么念，因为我不会说盖尔特语）即将交战。按照惯例，每个部落的首领挑选一名勇士，然后由这两名勇士进行决斗，获胜的一方有权决定统治部落的人选。“Garbh”部落的三个顶级勇士是Bevan（在盖尔特语中是指年轻的勇士），Cathal（战斗英雄）和Duer（英勇的）。“Conchubhair”部落的勇士分别是Fagan（易怒的），Guy（理智的）和Neal（冠军）。两个部落的首领对他们勇士了

解到得情况如下：与 Guy 和 Neal 对阵，Bevan 的获胜概率大于 Cathal；但与 Fagan 对阵，Cathal 的获胜概率大于 Bevan；与 Fagan, Guy 和 Neal 对阵，Cathal 的获胜概率大于 Duer；与 Bevan 对阵，Guy 最容易赢；与 Cathal 对阵，Neal 最容易赢；与 Duer 对阵，Fagan 最容易赢；与 Bevan 对阵，Fagan 的获胜概率大于 Neal；与 Cathal 对阵，Guy 的获胜概率大于 Fagan；与 Duer 对阵，Guy 和 Neal 获胜的概率相同。每个首领知道对方了解这些情况，每个首领知道对方知道对方了解这些情况等（也就是说，博弈是共同知识）。假设每个首领只想赢得决斗，你能推测出会选择谁参加决斗吗？

7. 根据图 3—30 所表述的策略式两人博弈，求出：

参与者2

参与者1		w	x	y	z
	a	1,2	0,5	2,2	4,0
	b	1,3	5,2	5,3	2,0
	c	2,3	4,0	3,3	6,2
	d	3,4	2,1	4,0	7,5

图 3—30

a. 重复剔除严格劣势策略后所剩下的策略。

b. 重复剔除弱劣策略后所剩下的策略。（这个过程的运作原理与重复剔除严格劣势策略相同，即剔除掉存在于每一步骤中的所有弱劣策略。）

8. 观察一下图 3—31 所描述的三人博弈情形。参与者 1 可以从 a_1,b_1 或 c_1 中选择任意一行，参与者 2 可以选择 a_2 或 b_2 中的任意一列。参与者 3 可以选择矩阵 a_3 或 b_3。表格中的第一个数字代表参与者 1 的收益，第二个数字代表参与者 2 的收益，最后一个数字是参与者 3 的收益。请写出重复剔除严格劣策略后所剩下的策略。

a_3

	a_2	b_2
a_1	3,1,0	2,3,1
b_1	0,3,1	1,1,0
c_1	1,0,2	1,2,1

b_3

	a_2	b_2
a_1	3,1,1	1,3,2
b_1	2,0,2	2,2,1
c_1	1,1,1	0,2,0

图 3—31

9. 一个团伙控制着北大街地区（位于马里兰大街和巴克利大街之间）的毒品交易。这个城市的布局如图 3—32 所示。该团伙头目设定毒品的销售价格，并指派两名成员到该地区贩卖。如果每名毒贩均可以将所收价款的 20% 作为自己的酬劳，那么就需要选择自己的去向：是北大街与马里兰大街的交叉口还是北大街与查尔斯大街、圣保罗大街、卡尔菲特大街亦或是巴克利大街的交叉口，后五个大街构成每个毒贩的策略组合。既然价格是由团伙头目设定，每个毒贩又只关心自己的利益，因而他们会选择能使自己的销售量最大化的地区。

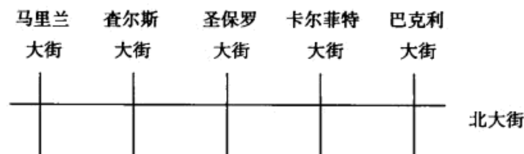


图 3—32

先简单假设五个大街之间的距离相等。顾客只住在北大街，而且平均分布在马里兰大街和巴克利大街之间（这样就不会出现顾客住在马里兰大街左侧或巴克利大街右侧的情况）。顾客知道这两个毒贩的卖价相同，所以他们会选择就近购买。北大街的毒品销售总量是固定的。唯一的问题是顾客是买毒贩子 1 的毒品还是毒贩子 2 的。这就意味着每名毒贩都想占据能够使自己的销售量最大化的地区。我们以每个毒贩的顾客数量作为他的收益代表（如图 3—33 所示）。通过分析当中的几项收益，我们就可以明白它们是如何被推断出的。譬如，假设毒贩 1 占据着马里兰大街和北大街的交叉口，毒贩 2 驻留在查尔斯大街和北大街的交叉口。所有住在马里兰大街和查尔斯大街之间的顾客购买毒贩 1 的毒品（因为他离他们最近），所有住在圣保罗大街右侧的顾客购买毒贩 2 的毒品。因而，毒贩 1 的市场占有率是 25%，毒贩 2 的则是 75%。由此可见，（$\frac{1}{4}$，$\frac{3}{4}$）是的策略组合（马里兰大街，圣保罗大街）的收益。现在假定毒贩 2 位于查尔斯大街，毒贩 1 位于马里兰大街。居住在马里兰大街和查尔斯大街之间的某个顾客觉得从谁那里买毒品无所谓。他左侧的所有顾客愿意从马里兰大街的毒贩那里购买毒品，并且这些顾客占该街区所有顾客数量的八分之一，那么策略组合（马里兰大街，查尔斯大街）的收益是（$\frac{1}{8}$，$\frac{7}{8}$）。设想两名毒贩都位于同一街口，如果他们各占有该地区一半的顾客，则收益是（$\frac{1}{2}$，$\frac{1}{2}$）。在重复剔除严格劣策略之后，推断出这两名毒贩的确切位置。

毒贩2的位置

毒贩1的位置	马里兰	查尔斯	圣保罗	卡尔菲特	巴克利
马里兰	$\frac{1}{2},\frac{1}{2}$	$\frac{1}{8},\frac{7}{8}$	$\frac{1}{4},\frac{3}{4}$	$\frac{3}{8},\frac{5}{8}$	$\frac{1}{2},\frac{1}{2}$
查尔斯	$\frac{7}{8},\frac{1}{8}$	$\frac{1}{2},\frac{1}{2}$	$\frac{3}{8},\frac{5}{8}$	$\frac{1}{2},\frac{1}{2}$	$\frac{5}{8},\frac{3}{8}$
圣保罗	$\frac{3}{4},\frac{1}{4}$	$\frac{5}{8},\frac{3}{8}$	$\frac{1}{2},\frac{1}{2}$	$\frac{5}{8},\frac{3}{8}$	$\frac{3}{4},\frac{1}{4}$
卡尔菲特	$\frac{5}{8},\frac{3}{8}$	$\frac{1}{2},\frac{1}{2}$	$\frac{3}{8},\frac{5}{8}$	$\frac{1}{2},\frac{1}{2}$	$\frac{7}{8},\frac{1}{8}$
巴克利	$\frac{1}{2},\frac{1}{2}$	$\frac{3}{8},\frac{5}{8}$	$\frac{1}{4},\frac{3}{4}$	$\frac{1}{8},\frac{7}{8}$	$\frac{1}{2},\frac{1}{2}$

图 3—33　毒贩在不同地理位置的收益

10. 两名学生即将考试，教授对他们说成绩较高者会得到 A，成绩较低者会得到 B。学生 1 的成绩等于 $x_1+1.5$，其中 x_1 是指她努力学习的程度（努力越多，成绩越高）。学生 2 的成绩等于 x_2，当中 x_2 代表她付出的努力。一个隐含的假设是学生 A 更聪明，如果付出的努力不变，学生 1 会凭借多出的 1.5 分获得较高的成绩。假定 x_1 和 x_2 可以在｛0，1，2，3，4，5｝之间任意取值，则学生 i 得 A 时得到的收益是 $10-x_i$，得 B 时得到的收益是 $8-x_i$，其中 $i=1$，2。

a. 重复剔除严格劣势策略后所剩下的策略。

b. 重复剔除弱劣策略后所剩下的策略。（这个过程的运作原理与重复剔除严格劣势策略相同，即剔除掉存在于每一步骤中的所有弱劣策略。）

11. 格罗克 • 马克思曾经说过："我永远也不会加入任何一个邀请我加入的俱乐部。"格罗克对加入你的投资俱乐部不感兴趣，但朱莉感兴趣。你的俱乐部现有十名成员，接纳新成员的程序很简单：每个人有接纳朱莉和不接纳朱莉两种选择的投票权，选择其中之一参加投票或者不参加投票选择弃权。朱莉必须收到六张赞同票才能够被接纳。设 m 为赞同票，收益函数如下：

$$\begin{cases}1 & \text{如果}m=6,7,8,9,10\\ 0 & \text{如果}m=0,1,2,3,4,5\end{cases}$$

a. 证明投赞同票（接纳朱莉）不是占优策略。

b. 证明选择弃权是弱劣策略。

c. 假设你劳累了一天，不想去参加定于今晚的投票，而不参加就意味着弃权，这一点可以由以下的收益函数列表看出：

$$\begin{cases}1 & \text{如果}m=6,7,8,9,10\text{并且你弃权}\\ \frac{1}{2} & \text{如果}m=0,1,2,3,4,5\text{并且你投票}\\ 0 & \text{如果}m=0,1,2,3,4,5\text{并且你弃权}\\ -\frac{1}{2} & \text{如果}m=0,1,2,3,4,5\text{并且你投票}\end{cases}$$

证明弃权不是弱劣策略。

12. 求出图 3—34 所示博弈的理性化策略。

参与者2

参与者1	x	y	z
a	0,4	1,1	2,3
b	1,1	2,2	0,0
c	3,2	0,0	1,4

图 3—34

附录 1

严格占优策略和弱占优策略

某博弈有 n 个参与者(1,2,……,n)。S_i 表示参与者 i 的策略组合，$s'_i \in S_i$ 的含义是策略 s'_i 是策略集 S_i 中的要素。S_{-i} 是 n-1 个参与者（不是参与者 i）的策略（n-1）元组的集合。$V_i(s'_i, s'_{-i})$ 是指当参与者 1 选择策略 s'_i 而其他参与者选择策略 s'_{-i}（$=s'_1,\cdots,s'_{i-1},s'_{i+1},\cdots,s'_n$）时他所得到的收益。由此，我们可以得出如下的定义：

1. 当且仅当 $V_i(s''_i, s_{-i}) > V_i(s'_i, s_{-i})$ 对所有的 $s_i \in S_{-i}$ 都成立时，策略 s''_i 严格优于策略 s'_i。也就是说，不管其他 n-1 个参与者选择什么策略，策略 s''_i 产生的收益都严格高于策略 s'_i。

2. 当且仅当 $V_i(s''_i, s_{-i}) > V_i(s_i, s_{-i})$ 对所有的 $s_{-i} \in S_{-i}$ 和 $s_i \neq s''_i$ 都成立时，策略 s''_i 是占优策略，即策略 s''_i 严格优于参与者的其他任何一个策略。
3. 当且仅当 $V_i(s''_i, s_{-i}) \geq V_i(s'_i, s_{-i})$ 对所有的 $s_{-i} \in S_{-i}$ 都成立，而且 $V_i(s''_i, s_{-i}) > V_i(s'_i, s_{-i})$ 对一些 $s_{-i} \in S_{-i}$ 成立时，策略 s''_i 弱优于策略 s'_i。换言之，对于其他参与者的所有策略，策略 s''_i 的收益和策略 s'_i 的一样多；对于其他参与者的一些策略，策略 s''_i 的收益严格高于策略 s' 的收益。

附录 2

可理性化策略

本章开始介绍说理性参与者不使用严格劣势策略。既然理性是指选择最佳策略，预测其他参与者的策略时，严格劣势策略的存在意味着不论其他参与者的策略是什么，都会有另一个产生严格更高收益的策略，因而，推测理性参与者不会选择严格劣势策略是合乎逻辑的。此外，假设所有参与者都相信所有参与者是理性的，那么每个参与者相信没有人会选择严格劣势策略。这样的逻辑认识引导我们剔除原有博弈中的严格劣势策略之后，继续剔除精炼后博弈中的严格劣势策略。如此循环往复，我们就推导出重复剔除严格劣势策略后所剩的策略。

如前所述，如果理性是共同知识，重复剔除严格劣势策略准则会剔除参与者不会选择的策略。但是他们会选择哪些策略呢？假设重复剔除严格劣势策略后，每个参与者只剩下一个策略，那么我们就可以清楚明确地预测出参与者的行为动向。一旦剔除掉所有看似不可能的因素后，剩下的无论多么荒谬都是事实。但如果重复剔除所有严格劣势策略后还剩下多个策略，该如何应对？虽然我们剔除了与理性是共同知识不符的策略，但剩下的策略就都符合理性是共同知识吗？

回想一下，理性表现为根据对其他参与者行为的信念选择最佳策略。因而，只有当至少一些对其他参与者策略的信念使该策略成为最佳策略时，这一策略才符合理性。我们直接运用这一定义，看看会发生什么。

以二人博弈为例。若参与者 1 是理性的，那么她会基于对参与者 2 策略的信念，选择使收益最大化的策略。但参与者 1 对参与者 2 策略的合理信念是什么呢？如果参与者 1 相信参与者 2 是理性的，那么基于她对参与者 2 所选策略的信念，她推测参与者 2 会选择收益最大化的策略。我们能允许参与者 1 根据对参与者 2 行为已有的信念来预测对方的策略选择吗？如果理性是共同知识，回答当然是否定的。假设参与者 1 相信参与者 2 相信参与者 1 是理性的，则参与者 1 相信参与者 2 对参与者 1 策略的信念应该是符合参与者 2 认为参与者 1 是理性的信念的。也就是说，参与者 2 相信参与者 1 选择的策略对于参与者 2 策略的一些信念是最佳策略。当然，分析还没有真正结束，但因为我有点头疼，所以还是进行概括性的陈述好了。

这个定义运用起来并不方便，所以用概述的方式来解释可理性化可能会产生事与愿违的效果。鉴于此，我们可以通过分析某个案例从另一方向着手解决。仔细观察图 3—35 所表述的博弈。

思考一下对于参与者 1 而言，策略 a 是否是可理性化的。存在这样一个关于参与者 2 策略的信念吗，它能够使策略 a 成为参与者 1 的最佳策略？当然存在，因为当且仅当参与者 1 相信参与者 2 使用策略 x 时，策略 a 就是最佳策略。参与者 2 有关于参与者 1 策略的信念吗，从而使策略 x 成为参与者 2 的最佳策略？假如没有的话，那么参与者 1 相信参与者 2 会使用策略 x 的信念将毫无意义（因为参与者 1 相信参与者 2 是理性的）。不存在这样的信念，也就没有必要讨论参与者 1 使用策略 a 的情况。实际上，当且仅当参与者 2 相信参与者 1 会使用策略 b 时，策略 x 是参与者 2 的最佳策略。

定义 3.4

可理性化（rationalizable）：若某策略符合理性是共同知识，则该策略是可理性化的。也就是说，如果信念本身符合理性是共同知识，那么该策略是某一参与者的最佳策略。

		参与者2		
		x	y	z
参与者1	a	3,1	1,2	1,3
	b	1,2	0,1	2,0
	c	2,0	3,1	5,0
	d	1,1	4,2	3,3

图 3—35　用可理性化策略求博弈的解

小结如下：若参与者 1 相信参与者 2 会使用策略 *x*，那么参与者 1 使用策略 *a* 是符合逻辑的。若参与者 2 相信参与者 1 会使用策略 *b*，则参与者 1 相信参与者 2 会使用策略 *x* 才有意义。然而，这又引发另一个疑问：参与者 2 相信参与者 1 会使用策略 *b* 是合理的吗？如果策略 *b* 是参与者 1 的严格劣势策略，那么相信参与者 2 使用策略 *x* 的信念会被削弱，参与者 2 使用策略 *a* 的论证也是如此。事实上，对参与者 1 而言，策略 *b* 严格劣于策略 *c*。假如参与者 1 相信参与者 2 相信参与者 1 是理性的，则参与者 1 不相信参与者 2 相信参与者 1 会使用策略 *b*，那么我们将无法使参与者 2 选择策略 *x* 理性化，因而也无法理性化参与者 1 对策略 *a* 的选择。策略 *a* 不是可理性化的，因为根据信念本身符合理性是共同知识，它不是参与者 1 的最佳策略。

对于策略 *c*，我们认为它是可理性化的。如果参与者 1 相信参与者 2 会使用策略 *z*，那么策略 *c* 就是她的最佳策略。可是，根据对参与者 1 策略的一些信念，策略 *z* 是参与者 2 的最佳策略吗？我们认为要让参与者 1 相信参与者 2 会使用策略 *z*，情况就得如此，这是因为参与者 1 相信参与者 2 是理性的并且理性参与者只使用基于自身信念的最佳策略。假设参与者 2 相信参与者 1 会使用策略 *d*，则策略 *z* 是参与者 2 的最佳策略。可参与者 2 相信参与者会使用策略 *d* 是合理的吗？没错，是合理的。因为如果参与者 1 相信参与者 2 会使用策略 *y*，则策略 *d* 是她的最优策略。因而，参与者 2 相信参与者 1 相信参与者 2 会使用策略 *y* 时，参与者 1 相信参与者 2 相信参与者 1 会使用策略 *d* 是合理的（你可能会认为这似乎永不完结，但是坚持一下，让你的大脑多动一动即可）。假使参与者 2 相信参与者 1 会使用策略 *c*，那么策略 *y* 是参与者 2 的最佳策略。这样一来，参与者 1 相信参与者 2 相信参与者 1 相信参与者 2 相信参与者 1 会使用策略 *c* 时，参与者 1 相信参与者 2 相信参与者 1 相信参与者 2 会使用策略 *y* 是合理的。好的，接下来我们从参与者 1 使用策略 *c* 开始，这个推断过程可以无穷尽地进行下去：

1. 参与者 1 相信参与者 2 会使用策略 *z* 时，策略 *c* 是参与者 1 的最佳策略。
2. 参与者 2 相信参与者 1 会使用策略 *d* 时，策略 *z* 是参与者 2 的最佳策略。
3. 参与者 1 相信参与者 2 会使用策略 *y* 时，策略 *d* 是参与者 1 的最佳策略。
4. 参与者 2 相信参与者 1 会使用策略 *c* 时，策略 *y* 是参与者 2 的最佳策略。
5. 参与者 1 相信参与者 2 会使用策略 *z* 时，策略 *c* 是参与者 1 的最佳策略。
6. 重复步骤 2 ~ 5。

多次“动动灰色脑细胞”（著名侦探赫尔克里•波洛之语）之后，我们推测出参与者 1 使用策略 *c* 是可理性化的，因为根据对参与者 2 策略的信念并且这些信念符合理性是共同知识，策略 *c* 是参与者 1 的最佳策略。而且，该循环中运用这些信念的所有策略都是可理性化的。例如，若参与者 2 相信参与者 1 会使用策略 *d*，则策略 *z* 是参与者 2 的最佳策略；若参与者 1 相信参与者 2 会使用策略 *y*，则策略 *d* 是参与者 1 的最佳策略；若参与者 2 相信参与者 1 会使用策略 *c*，则策略 *y* 是参与者 2 的最佳策略；若参与者 1 相信参与者 2 会使用策略 *z*，则策略 *c* 是参与者 1 的最佳策略，由此我们进行新的循环。所以，参与者 1 使用策略 *c* 和 *d* 是可理性化的，参与者 2 使用策略 *y* 和 *z* 是可理

性化的，而且这些策略也是唯一可理性化的策略。

假如你将重复剔除严格劣势策略准则运用于该博弈，会发现重复剔除严格劣势策略后所剩的策略正是刚刚推断出的可理性化的策略。很有趣吧！是巧合吗？不完全是这样。首先，可理性化的策略在重复剔除严格劣势策略之后能够继续存在是因为理性暗含着不使用严格劣势策略。可是，重复剔除严格劣势策略后所剩的策略中存在不可理性化的策略吗？是有可能的，尽管从专业的角度区分这两者的不同并不是我们在本书中所关注的。其次，在宏观的境况下，这两个概念可以推导出相同的答案。不难想象，重复剔除严格劣势策略准则很容易理解和运用，因而成为了我们关注的重点。但是，应当牢记的是可理性化的概念完全囊括了某一策略符合理性是共同知识所具有的含义。



第4章 稳定博弈：两人或三人非连续博弈中的纳什均衡

学习目标

4.1 纳什均衡的定义
4.2 经典的两人博弈
4.3 最优反应法
4.4 三人博弈
4.5 纳什均衡理论基础

本章导读

思想压倒一切的时代已经来临。

——维克多·雨果

起初，为了摆脱多部动画巨制撞档产生的麻烦，梦工厂早在2002年6月就宣称动画电影《鲨鱼黑帮》（*Sharkslayer*）的发行日期为2004年11月5日……梦工厂保留11月的发行日期被看成扇了迪士尼一记耳光，因为按照惯例，迪士尼会在该月发行它的皮克斯电影。迪士尼采取边缘政策，对外称不仅《鲨鱼黑帮》，动画电影《超人特工队》（*the Incredibles*）也会在11月5日进行首发，梦工厂则采取回避的态度，将电影的发行日期调至10月1日。

尽管被称为斗鸡博弈，但这个博弈其实是十分危险的。该博弈的典型形式是两位赛车手（通常是男青年为了追求感官上的惊险刺激）以类似决斗的方式正面交锋。两辆车猛冲向对方时，赛车手必须以近乎疯狂的方式快速决定是转弯以避免冲撞还是坚持冲下去（寄希望于对方会转弯）。博弈的目的是尽可能不采用转弯的策略，虽然双方都坚持冲下去的选择会导致血肉金属一片模糊的结果。斗鸡博弈的情景并不鲜见，譬如电影制片人（就发行日期）之间的博弈，美国和苏联领导人（就核武器方面）的博弈。表4—1列举了现实和虚构情境中出现的其他形式的斗鸡博弈。

图4—1所示的是策略式斗鸡博弈。[①]因为两个参与者都没有一个严格劣势策略，所以采用重复剔除严格劣势策略（参见3.3.2小节）的方法无法帮助我们求出该博弈的解。但别灰心，博弈理论总是能够给我们提供更多求解博弈的技巧。

①你可能并不赞同参与者都选择“坚持冲撞”的策略组合，因为这样做意味着受伤甚至死亡。你可以这么认为，但是这丝毫不影响博弈的结果。只要参与者选择“坚持冲撞”策略的收益小于矩阵中其他策略的收益，那么我们对参与者行为的预测都会趋同。这一状况反映了这样一个事实：重要的是策略的收益等级，而不是它们的实际价值。——作者注

表 4—1　斗鸡博弈运用

类型	案例
牵引器	电影《浑身是劲》（*Footloose*）
恐吓者	电视剧《发展受阻》（*Arrested Development*）中的布特（Buster）和戈贝（Gob）
轮椅	电视剧《万岁》（*Banzai*）中坐着机动轮椅的两位老妇人
雪地机动车	两名成年男性驾驶雪地机动车正面冲撞后死亡，斗鸡博弈以平局告终。（引自网页）
电影发行日期	梦工厂和迪士尼-皮克斯
核武器	古巴导弹危机："随着核武器僵局的日趋明显，东西方政府开始采取杜勒斯（Dulles）先生所称的'边缘政策'的策略。据我所知，该策略源自一些颓废的年轻人玩的一种游戏——'斗鸡'"。①

司机2

		转弯	坚持冲撞
司机1	转弯	2,2	1,3
	坚持冲撞	3,1	0,0

图 4—1　斗鸡博弈

如果你读过《美丽心灵》（*A Beautiful Mind*）这部小说或看过据此改编的电影，那么相信你不会对那名患有精神分裂症的天才数学家约翰·纳什感到陌生。在普林斯顿大学完成的博士论文中，纳什博士对博弈理论做出了两方面的贡献，其中之一就是著名的纳什均衡理论。40 年后，他凭借该理论获得了诺贝尔经济学奖。

本章中我们要完成的任务如下：4.1 节介绍纳什均衡的定义，学习如何用这一概念求出博弈的解，从而认识这一概念对于如何更好地理解人类行为的重要性。从 4.2 节开始，我们先着手分析一些囊括了策略情境中出现的冲突和共同利益的简单两人博弈。为了讨论更复杂的博弈情形，在 4.3 节中引入最优反应概念以求出纳什均衡的解，然后在 4.4 节中分析该方法应用于三人博弈的情况。最后，我们在 4.5 节中会更深入地了解假设参与者按照纳什均衡的要求行事的含义。

4.1 纳什均衡的定义

在理解纳什均衡以及它如何成为博弈求解的恰当方法之前，我们先回顾上一章所讲的内容。在某一博弈情境中，一个参与者能够根据对其他参与者策略的信念选择使自己收益最大化的策略时，就被认为是理性的。但问题是一个参与者对另一个参与者的策略拥有什么样的信念才是合理的？第 3 章通过假设理性是所有参与者的共同知识推导出这些信念。例如，若参与者 2 有一个占优策略并且参与者 1 相信参与者 2 是理性的，则参与者 1 相信参与者 2 会使用她的占优策略。这样就可以求出参与者 1 对参与者 2 策略的信念。

纳什均衡的方法假设参与者是理性的，然后采用另一方法明确这些信念。纳什均衡要求每个参与者对其他参与者策略的信念是正确的。比如，参与者 1 推测参与者 2 使用的策略正是参与者 2 实际采用的策略。纳什均衡的定义由以下两部分组成：

1. **参与者是理性的。**每个参与者基于对其他参与者策略的信念，采用使自己的收益最大化的策略。
2. **信念是正确的。**每个参与者对其他参与者策略的信念符合事实。

①引自《常识与核战争》（*Common sense and nuclear warfare*），伯特兰·罗素——作者注

前提（1）比较简单，前提（2）则相对比较难理解一些，它要求参与者能够对其他参与者的行为作出准确的预测。在某些情况下，这么做是合理的；而在其他的情形下，也许就不合理。理性行为假设和正确信念假设的组合就是纳什均衡的定义。

定义 4.1

纳什均衡（Nash equilibrium）：若每个参与者根据对其他参与者策略的信念，采用最大化自身收益的策略，则这样的一个策略组合就是纳什均衡。

有 n 个参与者，就需要满足 n 个条件。满足这一前提的一个策略组合才是纳什均衡，即每个参与者所选策略一定是对其他参与者所选策略的最优反应。所以，所有的参与者都同时作出最优反应。如果偏离其中的任何一个前提，那么该策略组合就不是纳什均衡。与马蹄铁博弈不同，不是说只差一个就满足所有条件的情形就近似于纳什均衡，**纳什均衡必须满足全部条件，否则什么都不是。**

将纳什均衡作为一个博弈求解概念是因为它假设策略组合是稳定的（每个参与者满足于针对其他参与者的策略而选择的策略）。例如，一个策略组合之所以不是纳什均衡是因为参与者 3 所选策略不是对其他参与者所选策略的最优反应。我们期待参与者 3 在发现该策略不是最佳策略后能改变策略。相反，纳什均衡不属于这样的事后预测，因为参与者很满足于自己所选的策略。

为了更详尽地了解这一点，设想一个参与者重复博弈的场景。假设他们目前没有按照纳什均衡的原则采取行动，在经过一轮的博弈后，会有一个参与者发现自己所选的策略不是对其他参与者所选策略的最优反应。为了提高自己的收益，他会改变策略。相比之下，如果他们按纳什均衡的原则行事，那么他们在每一轮的博弈之后都会满足于自己的策略。由纳什均衡产生的行为被认为会持续很长时间，社会学家通常对研究持续行为很感兴趣（不一定是因为不稳定行为很乏味，而是因为该行为更难以解释）。

希望以上的解释能够使你相信纳什均衡是一个很有价值的博弈求解概念（如果没做到的话，请见谅），接下来我们将这一概念应用于斗鸡博弈中。假设有四个策略组合，然后推断每个策略组合是否是纳什均衡。

- （坚持冲撞，坚持冲撞）。司机 2 选择坚持冲撞的策略，则司机 1 选择转弯策略的收益是 1，选择坚持冲撞的收益是 0（如图 4—2 所示）。所以，司机 1 会选择转弯策略（并忍受着朋友们的嘲笑），而不是坚持冲撞策略（以知道是否有来生）。这样，坚持冲撞策略不是参与者 1 的最优反应，这就意味着参与者 1 的纳什均衡条件不满足。至此，我们推断策略组合（坚持冲撞，坚持冲撞）不是纳什均衡。（既然我们已经推断出该策略组合不是纳什均衡，那么显然，驾驶员 2 的坚持冲撞策略也不是最佳策略。）

		司机2	
		转弯	坚持冲撞
司机1	转弯	2,2	[1],3
	坚持冲撞	3,1	[0],0

图4—2　斗鸡博弈：司机2选择坚持冲撞策略增加司机1的收益

- （转弯，转弯）。司机 1 选择转弯策略的话，那么司机 2 选取转弯策略的收益是 2，选取坚持冲撞策略的收益是 3（如图 4—3 所示）。如果司机 1 准备临阵退缩，那么司机 2 会坚持冲下去。因为转弯策略不是司机 2 的最佳策略，所以策略组合（转弯，转弯）也不是纳什均衡。

		司机2	
		转弯	坚持冲撞
司机1	转弯	2,[2]	1,[3]
	坚持冲撞	3,1	0,0

图4—3　斗鸡博弈：司机1选择转弯策略增加司机2的收益

- （转弯，坚持冲撞）。若司机 2 选择坚持冲撞策略，那么转弯策略是司机 1 的最优反应（该策略产生的收益 1 高于坚持冲撞策略的收益 0）。所以，转弯策略满足司机 1 所选策

略是最佳策略的条件。对于司机 2 而言，坚持冲撞的策略是最优反应，因为它产生的收益 3 高于转弯策略的收益 2。司机 2 选取的策略是她的最佳策略的条件也满足。每个司机所选的策略是对另一个司机所选策略的最优反应，所以策略组合（转弯，坚持冲撞）是纳什均衡。

- （坚持冲撞，转弯）。按照类似于上一个策略组合的推理过程可以推断出，该策略组合也是纳什均衡。

总的来说，该博弈中有两个纳什均衡 策略组合（转弯，坚持冲撞）和（坚持冲撞，转弯）。二者都不会发生汽车冲撞，而且只有一个司机会转弯。然而，纳什均衡并没有告诉我们哪个司机会转弯。

进行斗鸡博弈的最好办法也许是剔除策略组合中的转弯策略，而且最重要的是让对方也认识到这一点。图 4—4 表述的是司机 1 从他的策略组合中剔除转弯策略后的博弈格局。这样，就只剩下一个纳什均衡：司机 1 选择坚持冲撞，司机 2 选择转弯。

司机2

		转弯	坚持冲撞
司机1	坚持冲撞	3,1	0,0

图4—4 司机1剔除转弯策略时的斗鸡博弈

20 年前的海军交战就运用了类似于图 4—4 所示的策略。让我们来听听两位当事人的无线电通话。

1："为避免碰撞，请你方调整航线，方位角北 15 度。"

2："建议贵方向南 15 度改变（你的）航线，以免碰撞。"

1："我是这艘美国海军战舰的船长，我再重申一遍，调整你们的航线！"

2："我再重复一遍，调整'你的'航向。"

1："这艘'企业'号航空母舰，是一艘大型的美国海军战舰。你方迅速调整航线。"

2："这里是灯塔，你看着办吧。"

4.2 经典的两人博弈

这一章的主要目的是让你了解纳什均衡的概念和纳什均衡的求解方法。先从一些简单的两人博弈，其中的每个参与者最多有三个策略的情况开始吧。正如我们所看到的，每个博弈可以有一个纳什均衡，多个纳什均衡，或者没有纳什均衡。第一种情况属于理想型，因为我们可以准确地预测出参与者的行为。第二种情况属于"富人的窘境"，因为我们无法做出清晰明确的预测。最后一种情况是不存在纳什均衡，那么（至少在这一点上）也就没有什么可讨论的。虽然本章中不会涉及不存在纳什均衡情况下的博弈求解问题，但我们会在第 7 章中就该问题进行广泛的讨论。

你或许正在思考有没有多重纳什均衡求解的简便法则。比如，第 3 章提出了求得与理性是共同知识相一致策略的模式，即重复剔除严格劣势策略准则。遗憾的是，纳什均衡求解并没有像这样的方法。对于有限博弈（即参与者的数量是有限的，每个参与者的策略也是有限的）而言，唯一通用的法则是**穷举搜索**（exhaustive search）。该法则要求逐一分析每一个策略组合，然后判断是否是纳什均衡。不过，我们还是会介绍一些进行穷举搜索的便捷方法。

多重纳什均衡求解的一个重要的概念是一个参与者的**最优反应**（best reply，或最优回应）。**对于其他参与者的每个策略组合，一个参与者的最优反应是使她的收益最大化的一个策略。**因此，一个参与者不是仅有一个最优反应，而是一个对其他参与者每个策略组合的最优反应。而且，对于其他参与者的每个策略组合而言，如果能产生最高收益的策略不止一个，那么就会有不止一个最优反应。

一个纳什均衡可以被理解为每个参与者的一个策略组合，这个策略组合确保这个参与者选取的策略是对其他参与者选取策略的一个最优反应。这些与定义 4.1 所提出的 n 个条件相同，我们只是采用了稍稍不同的表述方式而已。

定义 4.2

最优反应（best reply）：如果其他n-1个参与者选取策略（$s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n$），那么参与者$i$对策略（$s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n$）的一个最优反应是使他的收益最大化的一个策略。

案例研究

囚徒困境

斯大林执政时期，一名乐队指挥正在火车上看一个乐谱。两名克格勃（KGB）官员认为这个乐谱是一份密码，随即逮捕了这名乐队指挥，尽管乐队指挥辩称该乐谱只是柴可夫斯基创作的小提琴协奏曲。第二天，审讯员对这名指挥说："你最好坦白，因为我们已经抓住了你的同伙柴可夫斯基，而且他现在正在交代。"

囚徒困境（我们之前研究的歌剧《托斯卡》就与这一情形类似）是博弈理论中被讨论最多的例子。一个犯罪团伙的两名成员被警方抓获，他们分别被关在不同的屋子里接受警方的盘问。他们被告知：如果一人坦白，另一人不坦白，坦白者获释，而不坦白者将被关三年；如果两人都坦白，那么他们会被判两年的刑罚；如果两人都不坦白，每人将被关一年。假设每个参与者的刑期越短收益越高，则该博弈的策略式如图 4—5 所示。

		罪犯2	
		坦白	抗拒
罪犯1	坦白	2,2	4,1
	抗拒	1,4	3,3

图4—5　囚徒困境

这一博弈只有一个纳什均衡，即两个参与者都选择坦白。我们先证明策略组合（坦白，坦白）是一个纳什均衡。若罪犯 2 坦白，则罪犯 1 选择坦白策略的收益是 2，选择不坦白策略的收益是 1。这样就满足了确保参与者 1 所选策略是最优反应的条件。对于参与者 2 来说，如果参与者 1 坦白的话，则她也坦白获得的收益是 2，不坦白的收益是 1。所以，参与者 2 所选策略是最优反应的条件也满足。因而，策略组合（坦白，坦白）是一个纳什均衡。

需要进一步指出两点内容。首先，囚徒困境是典型的对称博弈。如果参与者有相同的策略组合，转换策略就会转换收益，那么两人博弈就是对称博弈。譬如，某一策略组合是（不坦白，坦白），则罪犯 1 的收益是 4，罪犯 2 的收益是 1。如果对调他们的策略使之成为(坦白，不坦白)，则罪犯 1 的收益是 1，罪犯 2 的收益是 4。信息对称的条件在某种程度上暗示着选择相同策略的参与者会获得相同的收益。

其次，对称博弈的一个重要特征是，如果一个对称策略组合（如"坦白，坦白"）是一个参与者的最佳策略，那它也是另一个参与者的最佳策略。若罪犯 1 的策略是对罪犯 2 策略的最优反应，则罪犯 2 的策略也是对罪犯 1 策略的最优反应。博弈和策略组合的对称性使得参与者的均衡条件是相同的。因而，对于一个对称博弈中的一个对称策略组合而言，要么所有的纳什均衡条件都成立，要么都不成立。了解了这一特性之后，很容易看出如果选择坦白是罪犯 1 对罪犯 2 坦白策略的最优反应，那么坦白策略也是罪犯 2 的最优反应。


点拨

就一个对称博弈中的一个对称策略组合而言，如果一个参与者的策略是最优反应，则所有参与者的策略都是最优反应。

注意坦白策略也是一个占优策略。无论其他罪犯的策略是什么，坦白策略产生的收益严格高于不坦白策略的收益。稍加思索就会发现如果一个参与者有一个占优策略，则一个纳什均衡会促使她使用这个策略。一个参与者的策略成为纳什均衡的一部分，该策略必须是对其他参与者所取策略的最佳策略。因为一个占优策略总是唯一最佳策略，所以它一定会在纳什均衡中被使用。由此可见，若所有的参与者都有占优策略（如在囚徒困境中），则该博弈就有一个使用这些占优策略的纳什均衡。因而，策略组合（坦白，坦白）是囚徒困境中唯一的纳什均衡。




点拨

若一个参与者有一个占优策略，则一个纳什均衡要求该参与者使用占优策略。若所有的参与者都有一个占优策略，则只有一个纳什均衡，其中的每个参与者都使用她或他的占优策略。

协调博弈——行驶规范

我们接下来要讨论的是协调博弈（其特征表现为参与者在调整自身行为方面有一个共同的利益）的一个例子。多数成年人每天面对的一个协调博弈是选择靠左行驶还是靠右行驶。当然，靠左行驶（比如在英国和日本）还是靠右行驶（例如在美国和智利）并不是十分重要的事情，只不过是我们遵从的一个标准而已。

图 4—6 表示的是两个司机之间的博弈。很容易证明该博弈有两个纳什均衡。一个是特尔玛和路易斯都选择靠左行驶，另一个是二者都选择靠右行驶。若路易斯靠左行驶，则特尔玛也靠左行驶的收益是 1，靠右行驶的收益是-1。所以，特尔玛选择靠左行驶的策略是对路易斯的靠左行驶策略的最优反应。同理，路易斯选择靠左行驶的策略是对特尔玛靠左行驶策略的最优反应。事实上，因为这是一个对称博弈，我可以引用“凭借对称性”（by symmetry）这个关键词语推断出路易斯选择靠左行驶的策略也是最佳策略。这样，策略组合（靠左，靠左）就是一个纳什博弈。类似的，策略组合（靠右，靠右）也是一个纳什博弈。

		路易斯	
		靠左行驶	靠右行驶
特尔玛	靠左行驶	1,1	−1,−1
	靠右行驶	−1,−1	1,1

图4—6　行驶规范博弈

相反，策略组合（靠左，靠右）不是一个纳什博弈。路易斯选择靠右行驶的话，特尔玛选择靠左行驶的收益是-1，而选择靠右行驶会产生更高的收益 1。可以直接推断出策略组合（靠右，靠左）也不是一个纳什均衡。

纳什均衡并没有告诉我们司机们就哪个标准达成一致，而只是表示出他们就某个标准达成一致。历史表明社会的确会制定一个统一的行驶规范，但究竟是靠左行驶还是靠右行驶会因时间和空间的不同而不同。据估计，有 75% 的道路是靠右行驶的。虽然今天人们因为法律的要求而遵守某一行驶规范，但规范早在被立法之前就已经存在了（事实上它远远早于汽车的出现）。通常，法律只是将自发形成的习惯法则编撰成法律条文而已。


谜题

假定一个美国司机和一个英国司机在中间地带相对开出，每个人从自己的国家驾车出发。以上这些都是共同知识。哪个人会靠哪边行驶？这会演变成一个斗鸡博弈吗？纳什均衡能做出准确预测吗？

协调和冲突博弈——打电话

在行驶规范博弈中，有两个纳什均衡，而参与者认为两者没什么区别，因为靠左行驶和靠右行驶都一样。现在设想一个情形有两个纳什均衡，但参与者会做出不同的排序。

科琳和温妮正在通电话聊天，但是电话突然中断。应该是科琳给温妮打电话，还是温妮给科琳打电话呢？科琳和温妮是参与者，每个人都有一个策略组合（打电话，等电话）。如果需要继续聊天的话，每个人都愿意打电话，但更希望对方这么做。若双方都回电话，那么对方的电话占线，彼此再联系不上。显然，如果双方都不回电话，那么彼此也再联系不上。该博弈的策略式如图 4—7 所示。①

		温妮	
		打电话	等电话
科琳	打电话	0,0	2,3
	等电话	3,2	1,1

图4—7　电话博弈

①该博弈和著名的“性别战争”博弈相同，只是情境更加中性罢了。“性别战争”博弈讲述的是一对恋人，男的喜欢看拳击比赛，女的喜欢看歌剧，但两人更愿意在一起。——作者注

PDG

策略组合（打电话，打电话）不是一个纳什均衡。以科琳为例，她选择等电话获得的收益会高于打电话获得的收益0，所以如果温妮选择等电话，那么科琳选择打电话就不是最佳策略。策略组合（等电话，等电话）也不是一个纳什均衡。因为如果科琳选择等电话，那么温妮选择等电话策略的收益是1，而选择打电话策略会获得更高的收益2，因而温妮所选策略不是对科琳所选策略的最优反应。

一个参与者打电话和另一个参与者等电话的策略组合是一个纳什均衡。思考一下策略组合（打电话，等电话），其中科琳选择打电话。假设温妮在等科琳的电话，科琳则愿意打电话。因为打电话产生的收益是2而等电话的收益是1。假如温妮期待科琳给她打电话，那么她愿意等科琳的电话。对温妮而言，选择等电话策略的收益是3，而选择打电话的收益则是0。

如果打电话博弈是对称博弈（事实也是如此），那么我可以由对称性推出策略组合（打电话，等电话）是一个纳什均衡，策略组合（等电话，打电话）也是纳什均衡（这样就变成了温妮打电话）。要想弄清楚其中的原因，需要先证明这是一个对称博弈。若策略组合是（打电话，等电话），则科琳的收益是2，温妮的收益是3。如果对换策略使组合变为（等电话，打电话），那么科琳的收益是3，温妮的收益是2，而且他们选择相同策略获得的收益也相同。对称性意味着温妮在策略组合（等电话，打电话）中的纳什均衡条件与科琳在策略组合（打电话，等电话）中的纳什均衡条件相同；科琳在策略组合（等电话，打电话）中的纳什均衡条件与温妮在策略组合（打电话，等电话）中的纳什均衡条件相同。这样，倘若在策略组合（打电话，等电话）中，科琳发现选择打电话是最佳策略，那么温妮会发现在策略组合（等电话，打电话）中，选择打电话是最佳策略。**均衡条件要么都满足要么都不满足。**同样，假如在策略组合（打电话，等电话）中，温妮发现等电话是最佳策略，则科琳会发现在策略组合（等电话，打电话）中，等电话是最佳策略。

总而言之，由于是对称博弈，所以要么策略组合（等电话，打电话）和策略组合（打电话，等电话）都是纳什均衡，要么都不是。既然策略组合（打电话，等电话）是一个纳什均衡，那么按照对称性，策略组合（等电话，打电话）也是一个纳什均衡。

谜题

假如一个博弈是对称博弈，但均衡是不对称的，参与者如何协调？你如何在打电话博弈中协调？

猜谜博弈——石头、剪刀、布

莉萨：看来只有用猜石头、剪刀、布的办法解决了。

莉萨的想法：巴特的预测能力很差劲，总是选石头。

巴特的想法：出石头很不错，准赢。

（巴特出石头，莉萨出布）

巴特：唉！

——节选自《辛普森一家》

你曾经有多少次通过使用猜“石头、剪刀、布”的办法来解决分歧？假如你所在的文化不使用这样一个策略，那么下面就讲一下这种博弈。假设有两个人，每个人可以上下移动手臂四次。第四次，放下手时或握拳（表示出石头），或摊开手掌（表示出布），或伸出中指和食指（含义不言而喻）。获胜的规则如下：一人出石头，另一人出剪刀，则石头赢（因为剪刀剪不动石头）；一人出布，另一人出剪刀，则剪刀赢（因为剪刀能够剪开布）；一人出石头，另一人出布，则布赢（因为布能包住石头）。两人出拳相同的话，就是平局（多数情况下会继续猜直到产生赢家）。

设赢的收益是1，输的收益是-1，平局的收益是0，那么该博弈的策略式如图4—8所示。和巴特的信念相反，石头不是一个占优策略。虽然石头是对剪刀的最优反应，但不是对布的最优反应。事实上，该博弈中没有一个占优策略。每个策略是针对对手某个策略的最优反应。布是对石头的唯一最优反应，石头是对剪刀的唯一最优反应，剪刀是对布的唯一最优反应。

		莉萨		
		石头	布	剪刀
巴特	石头	0,0	−1,1	1,−1
	布	1,−1	0,0	−1,1
	剪刀	−1,1	1,−1	0,0

图4—8　“石头、剪刀、布”博弈

没有一个严格劣势策略，所以重复剔除严格劣势策略不能帮助我们走出困境（所有的策略都是经过重复剔除严格劣势策略后剩下的策略）。这时，优秀的博弈理论家会从工具箱里找出纳什均衡来帮助解决问题。经过一番努力之后，我们排除掉了一些策略组合。首先排除（石头，石头），因为如果莉萨出石头，巴特应该出布，那么，策略组合（石头，石头）就不是博弈的解（因为它不是一个纳什均衡）。接下来排除（布，石头）。巴特选择策略石头是一个最优反应，但莉萨选择策略石头却不是一个最优反应，因为巴特选择策略布时，莉萨选择剪刀的收益高于选择石头的收益。这样一来，策略组合（布,石头）就和策略组合（石头,石头）一起被否定掉了。我们继续乐此不疲地工作着，最后才发现一个糟糕的情况——所有的策略组合都被否定掉了。这九个策略组合中，竟然没有一个是纳什均衡。

在你逐一核对这九个策略组合以证实这个推断之前，我先介绍一个用来求解两人博弈的便捷之法。假设莉萨选择的某个策略（假定为 y）是一个纳什均衡的构成要素（之所以用构成要素这个词，是因为成为一个纳什均衡的条件还包括巴特所选的一个策略）。策略 y 要成为一个纳什均衡的构成要素，巴特选择的策略（假定为 c）必须是莉萨所选策略 y 的一个最优反应。只有选择这样的策略才能使巴特的纳什均衡条件成立。为了确保莉萨也会采用最佳策略，我们需要解出她对巴特所选策略 c（回想一下，该策略也是巴特对莉萨所选策略的最优反应）的最优反应。假使莉萨对策略 c 的最佳反应就是策略 y（这就是我们一开始为莉萨设定的策略），那么我们已经证明策略 y 的确是一个纳什均衡的构成要素，而事实上，这个纳什均衡就是策略组合（c, y）。但如果莉萨对巴特所选策略 c 的最优反应不是策略 y，那么我们推断策略 y 不是任何一个纳什均衡的构成要素。这样的话，我们一下子就剔除了所有涉及莉萨选择策略 y 的策略组合。以下就是假设莉萨选择策略 y 成为一个纳什均衡构成要素的结果：

> 莉萨选择策略 y→巴特对策略 y 的最优反应是策略 c→莉萨对策略 c 的最优反应是策略 y。

我们运用该法则来分析一下莉萨选择策略石头是否是一个纳什均衡的构成要素。如果巴特猜想莉萨会出石头，他就会出布。巴特选择出布的话，莉萨会选择出剪刀。因为这一选择和我们之前假定莉萨会选择出石头的策略不同，所以我们推断这不是莉萨选择出布的纳什均衡。因而，任何一个莉萨选择出石头的策略组合，即（石头，石头），（布，石头）和（剪刀，石头）都不是纳什均衡。现在用同样的法则来分析一下莉萨选择出布的情形，从而推断出她选择该策略是否是一个纳什均衡。莉萨选择出布，则巴特的最优反应是选择出剪刀，而莉萨对巴特选择出剪刀策略的最优反应是出石头，而不是布。因此，莉萨选择出布的策略也不是一个纳什均衡，从而得出所有的策略组合（石头，布），（布，布）和（剪刀，布）都不是纳什均衡的结论。同理，莉萨选择出剪刀的策略也不是任何一个纳什均衡的构成要素。最后，我们证明了“石头、剪刀、布”博弈中没有纳什均衡。

“石头、剪刀、布”博弈是猜谜博弈中的一种。在猜谜博弈中，最大化自己的收益需要你能够猜透对方（或其他参与者）的想法。也就是说，你会选择出其不意。对方猜测你选择策略 x 时，她会选择策略 b，而你为了应对她选择的策略 b 就不会选择策略 x，而是选择其他策略。比如，莉萨猜想巴特会出石头，所以她会出布；但巴特为了使莉萨猜不透他会出什么，会选择出剪刀，而不是石头。（很遗憾，巴特并不够聪明，但这只能归咎于马特•格罗宁 [Matt Groening] 而不是博弈理论。）

事实上，猜谜博弈会出现在许多的情形中。体育和军事竞赛就是其中的典型例子（我们会在第 7 章中对它们进行全面的分析）。不过，需要事先提醒一下的是，如果你打算加入美国“石头、剪刀、布”联盟（确实有这样一个组织），博弈论真的不能帮你获胜。

“石头、剪刀、布”不仅仅是儿童游戏的说法，近来它被两家顶级拍卖行——佳士得拍卖行和苏富比拍卖行之间的博弈所证实。某藏家决定选择一家拍卖行来拍卖自己的一件价值 2 000 多万美元的艺术品，被选中的

拍卖行（经过一轮“石头、剪刀、布”博弈产生）将会获得几百万美元的佣金。需要说明的是，该博弈的方式不是传统的举手表决，而是由佳士得拍卖行的一名经理和苏富比拍卖行的一名经理在一张纸上写出三个策略中的一个。结果，佳士得拍卖行因选择石头策略战胜了苏富比拍卖行的剪刀策略而中标。

◎案例分析

博弈中的冲突和共同利益

“石头、剪刀、布”是一个纯冲突博弈。这么说是什么意思呢？注意图4—8所示的收益矩阵有一个有趣的特性：参与者的收益之和总是等于同一个数（恰巧是0）。例如，巴特和莉萨都选择出石头，则每个人的收益是0，从而他们的收益总和是0。巴特选择出布，莉萨选择出石头，则巴特的收益是1，莉萨的收益是–1，将两人的收益相加还是0。这类博弈被称为常和博弈（constant–sum game），因为收益之和总是等于同一个数字。当该数字为0时，该博弈被称为**零和博弈**（zero–sum game）。

思考一下这其中的含义。收益之和总是等于同一个数字，所以如果巴特的某个策略组合产生较高收益的话，则莉萨的策略组合会产生较低的收益。因而，巴特的收益较高时意味着莉萨的收益较低；同理，莉萨的收益较高时则巴特的收益较低。从这一角度来说，“石头、剪刀、布”是一个纯冲突博弈。实际上，所有的常和博弈都有这一特性。

将该博弈与行驶规范博弈做一番对比的话会发现，后者正好与“石头、剪刀、布”博弈相反，即当中没有任何冲突。使司机1收益较高的策略同样可以给司机2带来较高的收益，比如策略组合（靠左，靠左）和策略组合（靠左，靠右），他们都会得到收益1而不是0。这就是**一个共同利益博弈，表现为一个参与者按照收益高低排序的策略组合恰好与另一个参与者的策略组合相同。**

斗鸡和电话博弈则刚好介于上述两种极端情况之间。这两个博弈的策略式情形使共同利益的存在成为可能。在斗鸡博弈中，两个参与者都不选择策略组合（坚持冲撞，坚持冲撞），而选择策略组合（转弯，坚持冲撞）和策略组合（坚持冲撞，转弯）。但也存在冲突的可能，因为他们对如何排列策略组合（转弯，坚持冲撞）和策略组合（坚持冲撞，转弯）意见不一，司机1倾向于选择后者，而司机2更可能会选择前者。同样，在电话博弈中，科琳和温妮都认为其中一人打电话要比双方等电话或双方打电话更可取，但是对究竟谁打电话持不同意见。科琳更想让温妮打电话，而温妮更希望科琳打电话。他们在协调需要一个人打电话方面存在共同利益，但在决定究竟由谁打电话时又发生了分歧，即存在冲突。

懂得分辨参与者的利益是完全冲突，部分冲突或是完全相同有助于我们推断哪些策略组合是纳什均衡。所以，当你分析某一博弈时，先不要机械地寻找解决方案，而是应该思考参与者之间所有的利益关系。反复思考可能会带给你一些有价值的简便方法。

4.1 理解测验

根据图4-9所示的博弈情形，求出所有的纳什均衡。

		参与者2			
		w	x	y	z
参与者1	a	0,1	0,2	1,0	3,2
	b	1,2	2,2	4,0	0,2
	c	2,1	0,1	1,2	1,0
	d	3,0	1,0	1,1	3,1

图4—9

4.3 最优反应法

把博弈人数增加至三人，就可达到著名烹饪节目主持人埃默里·拉加斯（Emeril Lagasse）常说的那句话“让我们更上一层楼”的境界。但在此之前，先介绍一个求解多重纳什均衡的简便方法。

回想一下，一个参与者的最优反应是针对其他参与者策略使自身收益最大化的一个策略。我们可以将一个纳什均衡看做一个策略组合。在这个策略组合中，每个参与者所选的策略都是对其他参与者所选策略的一个最优反应。从这一角度出发，最优反应法提供了一个推断所有纳什均衡的途径。与其做抽象化的描述，倒不如将该方法应用于两人博弈（见图4—10）中并进行说明更为有效。

		黛安		
		x	y	z
	a	1,1	2,1	2,0
杰克	b	2,3	0,2	2,1
	c	2,1	1,2	3,0

图4—10　杰克和黛安的策略

我们需要找出杰克对黛安所选的每个策略的最优反应。黛安选择策略 x，则杰克有两个最优反应——策略 b 和策略 c，当中每个策略的收益 2 都高于另一个可行策略 a 的收益 1。如果黛安选择策略 y，则杰克只有一个最优反应——策略 a；如果黛安选择策略 z，则策略 c 是杰克唯一的最优反应。为了便于记录下这些最优反应，我们会在所有与杰克的最优反应相关联的收益上画一个圈，如图 4—11 所示。

		黛安		
		x	y	z
	a	1,1	②,1	2,0
杰克	b	②,3	0,2	2,1
	c	②,1	1,2	③,0

图4—11　杰克的最优反应（画圈的）

下面，要找出黛安对杰克所选的每个策略的最优反应。如果杰克选择策略 a 的话，策略 x 和策略 y 都是黛安的最优反应；如果杰克选择策略 b，那么黛安的最优反应是策略 x；如果杰克选择策略 c，则策略 y 是黛安的最优反应。在所有与黛安的最优反应相关联的收益上打圈，如图 4—12 所示。

		黛安		
		x	y	z
	a	1,①	②,①	2,0
杰克	b	②,③	0,2	2,1
	c	②,1	1,②	③,0

图4—12　黛安和杰克的最优策略（画圈的）

因为纳什均衡是每个参与者的策略都是一个最优反应的策略组合，所以我们可以把图 4—12 中的纳什均衡看成是每个方格中两个收益都被打上圈的策略组合。这样，策略组合（b, x）和策略组合（a, y）都是纳什均衡。我们刚刚用最优反应法推导出所有的纳什均衡。在把最优反应法运用于三人博弈之前，我们先在“石头、剪刀、布”博弈中小试一把。图 4—13 表述的是标注出莉萨和巴特最优反应的情形。从图中可以看出，莉萨选择石头策略，则巴特的最优反应是选择布，所以我们将巴特获得收益 1 的策略组合（布，石头）圈住。注意没有一个方格有两个打圈的收益，这表明没有一个纳什均衡，而这一结果和我们之前的推论相同。

		莉萨		
		石头	布	剪刀
	石头	0,0	−1,①	①,−1
巴特	布	①,−1	0,0	−1,①
	剪刀	−1,①	①,−1	0,0

图4—13　莉萨和巴特在“石头、剪刀、布”博弈中的最优反应（画圈的）

4.4 三人博弈

◯ 案例研究

《美国偶像》迷博弈

艾丽西亚，凯特琳和劳伦斯是三个狂热粉丝。他们刚刚得到参加本周电视节目《美国偶像》的入场券。三个年轻人都支持剩余九个选手中的一个——埃斯·扬（Ace Young），他们决定利用这次机会来表明自己的态度。他们想出了一个计划：穿上写着“ACE”三个大字的 T 恤去参加节目。劳伦斯穿的 T 恤上面是一个大大的“A”，凯特琳的体恤上是“C”，艾丽西亚的是“E”。假如他们真这么做的话，天晓得会发生什么，他们可能会就此与全国电视节目无缘！天哪！

他们三个人都觉得这个主意很不错，但是他们也很想穿最近从 Bebe 淘来的新上衣。现在离他们出发去演播室碰面只剩下一个小时的时间了，可每个人都还在家里不知道是穿 Bebe 上衣还是穿标有字母的 T 恤。每个人应该穿什么呢？

要详细地说明该策略式博弈。先假定选择穿印有字母的 T 恤（很可能要记得按正确的顺序入座）得到的收益是 2。这一收益要高于选择穿 Bebe 上衣策略得到的收益 1。一个人穿印有字母的 T 恤，而其他两个人有一人或两人都不穿的话，则收益为 0。此时，这个身穿印着

字母T恤的女孩会发现结果真是糟糕透顶：她看起来一点也不迷人，而且他们也实现不了之前的愿望——表示出对埃斯·扬的支持。

图4—14表示的是该博弈的策略式。①劳伦斯的策略选择由行（穿印有字母A的T恤或穿Bebe上衣）表示，而凯特琳的策略选择和艾丽西亚的策略选择分别由列和矩阵表示。

艾丽西亚选择印有字母E的T恤（行：劳伦斯；列：凯特琳）

劳伦斯 \ 凯特琳	C	Bebe
A	2,2,2	0,1,0
Bebe	1,0,0	1,1,0

艾丽西亚选择Bebe上衣（行：劳伦斯；列：凯特琳）

劳伦斯 \ 凯特琳	C	Bebe
A	0,0,1	0,1,1
Bebe	1,0,1	1,1,1

图4—14 《美国偶像》迷博弈

运用最优反应法求出该博弈的解。对劳伦斯而言，如果艾丽西亚选择穿印有字母*E*的T恤，凯特琳选择穿印有字母*C*的T恤，则劳伦斯的最优反应是穿印有字母*A*的T恤。所以，我们圈住了方格中和策略组合（*A,C,E*）相关联的劳伦斯的收益2（如图4—15所示）。若凯特琳选择穿Bebe上衣，艾丽西亚选择穿标有字母E的T恤，则劳伦斯的最优反应是穿她的Bebe上衣并且得到收益1，我们将这个收益圈起来。艾丽西亚选择穿她的Bebe上衣，凯特琳选择穿印有字母*C*的T恤的话，劳伦斯的最优反应也是穿她的Bebe上衣，所以我们把她得到的收益1也圈住。其他两个女孩都选择穿他们的Bebe上衣，则劳伦斯的最优反应同样是穿她的Bebe上衣，这就意味着我们需要把她得到的收益1圈起来。凯特琳和艾丽西亚的情况可以用同样的方法来分析，至此，我们对图4—15的分析就结束了。

仔细观察图4—15，我们发现有两个纳什均衡，即两个策略组合，当中所有三个收益都被打上圈，表明三个少女都选择了最优反应。一个纳什均衡是劳伦斯选择穿印有字母A的T恤，凯特琳选择穿印有字母*C*的T恤和艾丽西亚选择穿印有字母*E*的T恤。这种情况下，看到有人在电视节目《美国偶像》中拼出自己的名字，埃斯·扬会非常高兴。另一个纳什均衡是每个人都选择穿上抢眼的Bebe上衣，而把标有字母的T恤抛之脑后。

艾丽西亚选择印有字母E的T恤（行：劳伦斯；列：凯特琳）

劳伦斯 \ 凯特琳	C	Bebe
A	②,②,②	0,1,0
Bebe	1,0,0	①,①,0

艾丽西亚选择Bebe上衣（行：劳伦斯；列：凯特琳）

劳伦斯 \ 凯特琳	C	Bebe
A	0,0,1	0,①,①
Bebe	①,0,①	①,①,①

图4—15 应用于美国偶像迷博弈中的最优反应法

注意该博弈中的参与者对纳什均衡的排序相同，即每个人都选择穿标有字母的T恤而不是Bebe上衣。这种情形不同于行驶规范博弈中参与者对纳什均衡排序持无所谓的态度，也不同于斗鸡和打电话博弈中每个参与者的纳什均衡排序结果各不相同的情形。

投票：真诚的还是虚假的

一个公司有三位股东。股东1掌握着25%的股份，股东2握有35%的股份，股东3持有40%的股份。该公司可以选择被另外两家公司*A*和*B*收购，也可以拒绝两家公司的收购请求。股东1按照从高到低的顺序对这三个选择进行了排序，结果如下：接受*A*公司的收购，接受*B*公司的收购和拒绝两家公司的收购（我们用选择*C*来表示）。股东2的选择倾向是*B*、*C*和*A*，股东3的选择倾向是*C*、*B*和*A*。表4—2所示的是三位股东选择倾向的排列等级。

表4—2 股东偏好排序

股东	首选	次选	末选
1	*A*	*B*	*C*
2	*B*	*C*	*A*
3	*C*	*B*	*A*

设一位股东的首选得到执行产生的收益是2，次选得以实施产生的收益是1，末选被执行产生的收益是0。三位股东同时投票。100张选票按照股东的股份进行分配，则股东1有25张选票，股东2有35张选票，股东3有40张选票。股东必须将选票全部投给某一个方案。例如，股东1只能将25张选票投向*A*、*B*或*C*当中的任

①该博弈是受法国思想家卢梭所著的《论人类不平等的起源》（*On the Origins and Foundations of Inequality among Men*）一书的启发，对猎鹿博弈进行改编的21世纪少女版本。在猎鹿博弈中，猎手们可以合作抓住一头鹿（而不是拼写出ACE），或者单独行动抓住一只兔子（不是穿一件Bebe上衣）。——作者注

意一个，而不能在这三个方案之间进行分摊。每个参与者的策略组合由 *A*、*B* 和 *C* 组成。得票数最多的方案将得到执行。

为了求出收益矩阵，首先需要推断投票如何产生多数票得主。例如，股东 1 和 2 赞成方案 *B*，则方案 *B* 会以 60 票（如果股东 3 赞成方案 *A* 或 *C*），或 100 票（如果股东 3 也赞成方案 *B*）获选。图 4—16 表述的是 27 种情形（三个参与者参加投票）中每种情形的多数票得主。

三人赞同方案*A*的情形

1 \ 2	*A*	*B*	*C*
A	*A*	*A*	*A*
B	*A*	*B*	*A*
C	*A*	*A*	*C*

三人赞同方案*B*的情形

1 \ 2	*A*	*B*	*C*
A	*A*	*B*	*B*
B	*B*	*B*	*B*
C	*B*	*B*	*C*

三人赞同方案*C*的情形

1 \ 2	*A*	*B*	*C*
A	*A*	*C*	*C*
B	*C*	*B*	*C*
C	*C*	*C*	*C*

图 4—16　多数票获得者

接下来用联合收益向量替换单元格中的各个备选方案。例如，方案 *B* 获选，则股东 1 的收益是 1（因为 *B* 是他的次选方案），股东 2 的收益是 2（因为 *B* 是他的首选方案），股东 3 的收益是 1（因为 *B* 是他的次选方案）。替换之后就产生了图 4—17。

描述这些股东如何进行投票时，会自然想到政治学家所说的真诚投票（sincere voting）的可能性。这个词是指投票人会按自己的首选情形进行投票。既然如此，那么股东 1 会把她的 25 张选票投给方案 *A*，股东 2 会将他的 35 张选票投给方案 *B*，股东 3 会把她的 40 张选票投给方案 *C*。所以，方案 *C* 会得到批准，因为它获得的票数最多。但真诚投票是一个纳什均衡吗？股东们选择真诚投票是最佳策略吗？事实上，不是最佳策略。注意股东 1 会选择方案 *B* 而不是方案 *C*。如果股东 2 和 3 选择真诚投票策略，那么股东 1 可能会进行虚假投票（如果可以这么命名的话），把选票投给方案 *B* 而不是方案 *A*。这样，方案 *B* 的选票是 60 票（得到了股东 1 和 2 的支持），成功获选。把她的选票从首选方案 *A* 转向她的次选方案 *B*，股东 1 的收益将会从 0 增加至 1。因此，真诚投票策略不是该博弈的一个纳什均衡。

三人赞同方案*A*的情形

1 \ 2	*A*	*B*	*C*
A	2,0,0	2,0,0	2,0,0
B	2,0,0	1,2,1	2,0,0
C	2,0,0	2,0,0	0,1,2

三人赞同方案*B*的情形

1 \ 2	*A*	*B*	*C*
A	2,0,0	1,2,1	1,2,1
B	1,2,1	1,2,1	1,2,1
C	1,2,1	1,2,1	0,1,2

三人赞同方案*C*的情形

1 \ 2	*A*	*B*	*C*
A	2,0,0	0,1,2	0,1,2
B	0,1,2	1,2,1	0,1,2
C	0,1,2	0,1,2	0,1,2

图 4—17　策略式投票博弈

虽然可以证明在只有两种选择的情况下，选择真诚投票是最佳策略，可是当存在三种或三种以上选择（正如我们刚刚所看到的）时，人们可能会放弃首选转而选择其他。直觉告诉我们出现这种情况是因为首选在现实中往往行不通，即不管你怎么投票，它都不会获选。既然首选不可行，那么就有理由考虑根据各种投票方案，

选择剩下的选项哪个是可行的，然后把票投给这当中的首选。在我们之前研究的例子中，股东2和3分别把票投给了方案*B*和*C*，这时，股东1可以让方案*B*获选（把她的票投给方案*B*），也可以让方案*C*赢（把票投给方案*A*或*C*），可问题是她究竟会选择方案*B*还是*C*。如果她选择方案*B*，那么她就应该战略性地进行投票以保证该方案获选。

证明真诚投票策略不是一个纳什均衡之后，需要分析用最优反应法是否可以解出一个纳什均衡策略组合。先从股东1开始。假设股东2和3赞同方案*A*，则不论股东1赞同方案*A*、*B*或*C*，她得到的收益都是2（该说法是有道理的，因为不管股东1怎么投票，方案*A*都会获得多数票），因而股东1的三个策略都是最优反应。在图4—18中，我们把与股东1和股东2选择方案*A*相关联的竖列中的收益2以及和股东3选择方案*A*相关联的矩阵中的收益都圈了起来。如果股东2赞同方案*B*，股东3赞同方案*A*，那么股东1的最优反应是赞同方案*A*或*B*（以确保方案*A*赢），相关的收益2被圈了起来。股东2赞同方案*C*，股东3赞同方案*A*，则股东1的最优反应是方案*A*和*C*。用这种方法继续分析股东1的情形后，着手分析股东2和3的情形，从而可以得出图4—18。

现在寻找三个收益都被画上圈的所有策略组合。在这样的一个策略组合中，每个股东所选的策略都是一个最优反应，因而不论其他股东做什么，每个股东都在尽自己最大的努力。也就是说，它是一个纳什均衡。细心观察图4—18，可以看出三个股东都使用最优反应的策略组合有五个，因而这五个策略组合——（*A*,*A*,*A*），（*B*,*B*,*B*），（*C*,*C*,*C*），（*A*,*C*,*C*）和（*B*,*B*,*C*）都是纳什均衡。需要注意的是这些纳什均衡会产生不同的结果：策略组合（*A*, *A*, *A*）会使方案*A*通过，因为大家都赞同方案*A*；策略组合（*B*, *B*, *B*）和（*B*, *B*, *C*）会使方案*B*通过；策略组合（*C*, *C*, *C*）和（*A*, *C*, *C*）会使方案*C*获选。

我们已经机械地推导出一组纳什均衡。尽管这具有一定的价值，但我们更需要明白它们是如何成为纳什均衡的。思考一下纳什均衡（*A*, *A*, *A*）。为什么股东2和3选择他们的末选方案是最佳策略呢？答案是这两名股东都不是关键要素，因为无论他们每个人如何投票，最终的结果都是一样的，即方案*A*获选。考虑一下股东2面对的情形。如果她赞同方案*A*，则方案*A*以100票获选；如果她赞同方案*B*，则方案*A*以65票获选；如果她赞同方案*C*，则方案*A*仍然以65票获选。股东2和3确实可以通过合作以得到更高的收益，因为如果他们都赞同方案*B*，那么方案*B*获选，他们也因此分别获得收益2和1，高于方案*B*获胜所得到的收益0。但是，参与者之间是不可以这样协调的。纳什均衡规定每个参与者只能在不与人沟通的条件下才能获得更高收益。①

三人赞同方案*A*的情形

2

1		A	B	C
	A	②,⓪,⓪	②,⓪,0	②,⓪,0
	B	②,0,0	1,②,①	②,0,0
	C	②,0,0	②,0,0	0,①,②

三人赞同方案B的情形

2

1		A	B	C
	A	②,0,⓪	①,②,0	①,②,1
	B	1,②,1	①,②,①	①,②,1
	C	1,②,1	①,②,1	0,1,②

三人赞同方案C的情形

2

1		A	B	C
	A	②,0,⓪	0,①,②	⓪,①,②
	B	0,1,②	①,②,①	⓪,1,②
	C	0,①,②	0,①,②	⓪,①,②

图4—18　运用于投票博弈的最优反应法

纳什均衡（*A*, *A*, *A*）的另一个有趣的特征是：股东2和3赞同方案*A*采用的是弱劣策略。如表4—3所示，股东2赞同方案*A*的策略弱劣于赞同方案*B*的策略。“股东2赞同方案*B*”竖列中的所有收益至少等于“股东2赞同方案*A*”竖列中的所有收益，甚至严格高于某些横排中的收益。所以，不论股东1和3如何投票，股东2

①注意，在这五个纳什均衡中的每一个均衡中，至少有一个股东不是关键要素。对于策略组合（*B*, *B*, *B*）和（*C*, *C*, *C*），三个股东都不是关键要素，就如同在策略组合（*A*, *A*, *A*）和（*A*, *C*, *C*）中，股东1不是关键要素一样，尽管股东2和3是关键要素，因为如果他们赞同方案*A*，则每个人都可以使方案*A*获胜。对于是策略组合（*B*, *B*, *C*）而言，虽然股东1和2是关键要素，但股东3不是。——作者注

赞同方案 *B* 策略的收益至少和赞同方案 *A* 策略的收益一样多。当然，股东 1 和 3 赞同方案 *A* 时，如同在纳什均衡（*A*, *A*, *A*）中的策略，股东 2 赞同方案 *A* 和 *B* 得到的收益都是 0，所以赞同方案 *A* 是她的最佳策略。然而，股东 2 和 3 还有其他的投票选择（例如，其中一人赞同方案 *A*，另一人赞同方案 *B*）。在这种情况下，股东 2 选择赞同方案 *B* 策略严格优于选择赞同方案 *A* 策略。

表 4—3　股东 2 的收益

股东1的策略	股东3的策略	两个股东赞同方案*B*		两个股东赞同方案*A*
A	*A*	0	=	0
A	*B*	2	>	0
A	*C*	1	=	0
B	*A*	2	>	0
B	*B*	2	=	2
B	*C*	2	>	1
C	*A*	0	=	0
C	*B*	2	=	2
C	*C*	1	=	1

这样，一个参与者选择一个弱劣策略的可能性就没有被纳什均衡排除。虽然股东 2 选择赞同方案 *B* 策略的收益至少和选择赞同方案 *A* 策略的收益相同（而且在某些情况下，会产生严格更高收益），但是对于股东 1 和股东 3 实际使用的策略而言，只要赞同方案 *A* 策略的收益和赞同方案 *B* 策略的收益一样多，那么方案 *A* 就是一个最优反应，因而也是一个纳什均衡。



点拨

一个纳什均衡不排除一个参与者使用弱劣策略的可能性。

晋级与阴谋破坏

假设你参加一项竞赛，成绩最高者将赢得奖金。目前，你排名第二。为了提高你获胜的概率，你会采取什么办法？一个可行的办法是努力学习从而提高比赛成绩。然而，更有效的办法可能是用卑劣的手段来降低暂时排名第一的对手的比赛成绩。你的目标是成为最终的冠军，所以你可以采取通过努力一步步爬上去的方式，也可以用把前面的对手拉下来的方式实现该目标。

这种具有破坏性的竞争模式常常出现在政治角逐中。下次总统预选开始时，关注一下这一竞选活动。落后的候选人不仅会谈论让他们当选总统是一个多么好的选择，也会谈到让领先的候选人当选是一个多么糟糕的选择。他们通常不会把时间浪费在诋毁其他候选人上，但暂时排名领先的那名候选人（也是他们要战胜的对手）除外。

为了探究卑劣手段是在何时以何种方式将领先者拉下台的，我们假定某一情形中有三个参与者争夺晋级权利，成绩最高者晋级。每个参赛者只能选择以下三种可能途径中的一种：要么努力提高自己的成绩（我们称之为积极努力），要么诋毁其他两名竞争者中的一名（我们称之为消极努力）。

比赛开始之前，假设参与者 *i* 的成绩是 v_i。如果一个参与者积极努力，则她的成绩增加 1。如果一个参与者选择诋毁参与者 *i* 的消极努力，则参与者 *i* 的成绩被减去 1。两个参与者都选择诋毁参与者 *i* 的消极努力，则参与者 *i* 的成绩被减去 4。因此，第二个参与者消极努力的边际影响要比第一个参与者的消极努力更具破坏力。这么讲似乎是合理的，因为一个人诋毁另一个人可能会被看做无稽之谈，但两个人这么做就会被认为是事实。

表 4—4 概述了努力是如何影响成绩的。譬如，参与者 *i* 选择积极努力，其他两个参与者选择诋毁她的消极努力，则参与者 *i* 最后的成绩是 v_i-3，即自身正面努力增加的 1 和被两人诋毁而被减去的 4 相加的结果。

表 4—4　参与者 *i* 的成绩

参与者*i*选择积极努力所增加的分值	参与者*i*被诋毁时所减去的分值	总成绩
0	0	v_i
0	1	v_i-1
0	2	v_i-4
1	0	v_i+1
1	1	v_i
1	2	v_i-3

假设参与者只在乎是否能晋级，而不关心比赛成绩。具体而言，就是一个参与者的收益取决于他晋级的可能性。如果一个参与者最后的成绩高于其他两个参与者，则她晋级的可能性是 1，所以她的收益是 1。如果她与另外一个参与者的成绩都是最高的，则她晋级的可能性

是$\frac{1}{2}$，因而每个人得到的收益都是$\frac{1}{2}$。如果三个参与者最后的成绩相同，每个人得到的收益是$\frac{1}{3}$。如果一个参与者的成绩是最低的，则她的收益是0，她晋级的可能性是0。

假定$v_1=2$且$v_2=0=v_3$时，参与者1领先。开始分析前，先理想化地认为不存在卑劣手段的策略组合，即每个参与者都选择积极努力，所以参与者i最后的成绩是v_i+1的情形。那么，参与者1最后的成绩就是3（因为她起初的成绩是2），参与者2和3最后的成绩是1（因为他们开始的成绩是0）。因此，参与者1晋级。如果所有的参与者都为了提高自己的成绩而选择积极努力，则开始领先的参与者最后的成绩也是最高的，从而成功晋级。我们现在分析一下这是否是一个纳什均衡。

- **参与者1**：显然，参与者1所选的策略是最佳策略，因为她的收益是1（即晋级的可能性），而且该收益是她能够得到的最高收益。因此，参与者1的其他策略不可能会产生更高的收益。
- **参与者2**：参与者2选择积极努力策略的收益是0，因为他的成绩1低于参与者1的成绩3，没有晋级的可能性。如果他选择诋毁参与者3的消极努力，结果也是于事无补，因为参与者2要打败的是参与者1，那么选择诋毁参与者3的消极努力并不会影响参与者1的成绩。最后一个可能性就是参与者2选择诋毁参与者1的策略，这样一来，参与者1的成绩就由3变为2，而参与者2的成绩也从1降至0（因为他不再选择积极努力的策略）。这种情况下，参与者2依旧晋级无望。由此看出，参与者2选择三个策略中的哪一个无所谓，因为所有策略的收益都是0。因而，参与者2愿意选择积极努力，因为没有一个策略能够产生一个严格更高收益。
- **参与者3**：参与者3的情形与参与者2的类似。他们的收益相同，选择的策略也相同。因此，如果参与者2选择积极努力的策略是最佳策略，那么积极努力的策略也是参与者3的最佳策略。
- 总之，三个参与者都选择积极努力是一个纳什均衡，因而成绩领先者会晋级。

在分析一个实施消极努力的策略组合时，我们考虑参与者的动机和一个合乎常理的策略组合。对于参与者2来说，选择诋毁参与者3的消极努力并没有什么意义，因为他要打败的人是参与者1（她的成绩暂时领先）。同理，参与者3也会这么做。参与者1应当把精力放在提高成绩上，因为她的成绩第一，因而她晋级的关键是继续保持领先。

同样，分析参与者1选择积极努力而其他参与者选择诋毁她的消极努力的一个策略组合，我们得到的结果如下：参与者1的成绩是−1（因为开始时她的成绩是2，选择积极努力会增加1，但其他两个参与者的消极努力会让她的成绩被扣掉4），参与者2和3的成绩是0（因为没有针对他们的积极或消极努力，所以他们的成绩和起初的一样）。参与者2和3同时获得最高成绩时，参与者1的收益是0，参与者2和3的收益都是$\frac{1}{2}$。下面看是否可以求出一个纳什均衡。

- **参与者1**：很遗憾，参与者1无法改变她的现状。倘若她选择诋毁参与者2的消极努力，那么参与者2的成绩会降至−1，她的成绩会减少至−2。参与者3会凭借她的成绩0而成功晋级，所以她仍然无法实现晋级的目的。以此类推，参与者1选择诋毁参与者3的消极努力会使得参与者2晋级，她同样达不到目的。所以，参与者1的最佳策略是选择积极努力。
- **参与者2**：如果参与者2选择诋毁参与者3（不是参与者1）的消极努力，那么参与者1的成绩会从−1增加至2，参与者2的成绩仍是0，参与者3的成绩会从0降至−1。既然参与者1晋级会导致参与者2的收益从$\frac{1}{2}$降至0，因而，参与者2会选择诋毁参与者1（而不是参与者3）的消极努力。参与者2选择积极努力会怎么样呢？这么做会使他的成绩增加至1，从而高于参与者3的成绩（仍然是0），但也会使参与者1的成绩增加至2（因

为只有一个人选择诋毁她的消极努力）。参与者 1 的成绩最高，所以参与者 2 的收益又是 0。于是，参与者 2 选择诋毁参与者 1 的消极努力策略严格优于选择诋毁参与者 3 的消极努力策略或选择积极努力的策略。

- **参与者 3**：按照分析参与者 2 的方式可以得出，参与者 3 选择诋毁参与者 1 的消极努力策略是最佳策略。
- 简而言之，策略组合——参与者 1 选择积极努力和参与者 2、3 选择诋毁参与者 1 的消极努力，是一个纳什均衡。这种做法可以降低参与者 1 的成绩，从而使得参与者 2 或参与者 3 晋级。成绩领先者失败了。正如韦恩•坎贝尔（Wayne Campbell）在节目《反斗智多星》（*Wayne's World*）中常说的："晋级的机会……被剥夺！"

晋级博弈中有多个纳什均衡（实际上，要比我们讲到的多），并且具有不同的内涵。一个纳什均衡中所有的参与者都努力提高自己的成绩，其结果印证了"最佳人选获胜"这句话。但在另一个较为阴险的解决方案中，最佳人选在其他两个实力较弱的参与者的联合对付之下被踢出了比赛。晋级成为了实力较弱两人中其中一人的囊中之物。这一结果用篮球运动员和经理人迪奥•迪罗谢（Leo Durocher）的话"失败的总是好人"来表述或许更为恰当。

4.2 理解测验

根据图4—19所示的博弈情形，写出所有的纳什均衡。

参与者3：*I*

参与者1 \ 参与者2	*x*	*y*	*z*
a	2,1,2	0,0,2	1,2,3
b	0,3,1	2,2,4	3,1,0
c	1,1,1	3,2,1	2,2,2

参与者3：*II*

参与者1 \ 参与者2	*x*	*y*	*z*
a	2,1,3	1,0,3	1,0,4
b	1,2,1	3,3,3	1,1,1
c	1,2,1	1,0,0	2,1,2

图 4—19

4.5 纳什均衡理论基础

到目前为止，我们已经介绍了博弈求解的两种方法，分别是重复剔除严格劣势策略和纳什均衡。但这两者有什么关联呢？接下来我们将回答这个问题，从而也引出了在纳什均衡情境中策略是如何被阐释的问题。事实证明，一个策略具有双重作用。

4.5.1　与理性是共同知识的关系

在探究重复剔除严格劣势策略和纳什均衡中的策略之间具有什么样的关系之前，我们先举一个例子。假定一个三人博弈纳什均衡，其中参与者 1 使用策略 x，参与者 2 使用策略 y，参与者 3 使用策略 z。这些策略能留存于重复剔除严格劣势策略中吗？很容易推断出第一轮剔除之后它们仍然存在，因为策略 x 是对参与者 2 所选策略 y 和参与者 3 所选策略 z 的一个最优反应，策略 x 不是一个严格劣势策略。同理，因为策略 y 是参与者 2 对参与者 1 所选策略 x 和参与者 3 所选策略 z 的一个最优反应，所以策略 y 不是一个严格劣势策略。最后，因为策略 z 是参与者 3 对参与者 1 所选策略 x 和参与者 2 所选策略 y 的一个最优反应，因而策略 z 不是一个严格劣势策略。因此，策略 x、y、z 会存在于第一轮的重复剔除严格劣势策略之中。

第二轮的情况如何呢？尽管参与者 2 和 3 的一些策略在第一轮中被剔除，但是策略 y 和 z 没有被剔除，所以策略 x 不是一个严格劣势策略。同理，在第二轮中，策略 y 不是参与者 2 的一个严格劣势策略，策略 z 也不是参与者 3 的一个严格劣势策略。这样，策略 x、y、z 继续存在于第二轮当中。就像那只劲量兔子（美国电池广告中永远动不停的粉色兔子）一样，这一推论会一直进行下去，对每一轮都适用。所以，如果策略组合 (x, y, z) 是一个纳什均衡，那么这些策略就能留存于重复剔除严格劣势策略中。虽然我们仅证明了三人博弈具有这样的特性，但实际上这一特性具有普适性，因而可以应用于所有的博弈中。

既然每个纳什均衡都留存于重复剔除严格劣势策略之中，那么是否存在一个策略留存于重复剔除严格劣势策略之中，却不是一个纳什均衡的构成要素的情况呢？

完全存在这样的可能性，事实上本章就有很多这样的例子。在美国偶像迷博弈中，所有的策略都留存于重复剔除严格劣势策略之中，因为没有一个策略是严格劣势策略。所以，重复剔除严格劣势策略准则允许这八个可行的策略组合中的任何一个策略存在下去。相反，只有（A, C, E）和（B, B, B）两个策略组合是纳什均衡。另外一个例子是"石头、剪刀、布"博弈。该博弈中所有的策略组合都留存于重复剔除严格劣势策略之中，但却没有一个是纳什均衡。纳什均衡要比重复剔除严格劣势策略更为严密，因为满足纳什均衡条件的策略组合数量更少。

点拨

所有纳什均衡都满足重复剔除严格劣势策略的条件，因而与理性是共同知识相一致。然而，留存于重复剔除严格劣势策略之中的一个策略组合不一定是一个纳什均衡。

图 4—20 表述的是纳什均衡如何成为留存于重复剔除严格劣势策略之中的策略组合的一个子集，而重复剔除严格劣势策略本身则是所有策略组合的一个子集。可是，对于某一特定博弈，这些集合都可能会重叠。比如，一个纳什均衡集合可能与留存于重复剔除严格劣势策略之中的策略组合相同，或者留存于重复剔除严格劣势策略之中的策略也许会重叠于由所有组合组成的集合。



图 4—20　纳什均衡和重复剔除严格劣势策略后所剩策略之间的关系

4.5.2 重新探讨策略的定义

为了更好地理解一个策略在纳什均衡情形中的作用，思考一下如何明确参与者 i 的一个策略（用 s_i 表示并代表他的决策规则）和参与者 i 对参与者 j 所选策略（用 $s_j(i)$ 表示）的猜测（表示的是参与者 i 对参与者 j 会采取什么方式博弈的判断）。一个策略组合（$s'_1,\dots s'_n$）是一个纳什均衡，如果对于所有的参与者 i 都有：

1. 当 $j\neq i$ 时，根据他对参与者 j 选择 $s_j(i)$ 的信念，s'_i 最大化参与者 i 的收益。
2. 当 $j\neq i$ 时，$s_j(i)=s_j'$。

策略 s'_i 在一个纳什均衡中起着双重作用。正如条件 1 所示，它是参与者 i 的决策规则。此外，按照条件 2 的描述，s'_i 是参与者 j 对参与者 i 会采取什么方式博弈的（准确）判断。

在 2.2 小节中，我们规定一个策略必须在每个可行的信息集合中明确一个参与者的行为。也就是说，一个策略必须在甚至无法得到一个信息集合的条件下根据之前某个信息集合所规定的行为做出行为判断。例如，在绑架博弈中，绑架者的策略是明确选择释放还是杀掉受害人，虽然在起初的节点该策略规定他不要实施绑架。一个策略必须满足这个必要条件，因为一个均衡策略具有双重作用。尽管一个参与者不会像预测的那样行事，但另一个参与者仍会猜测他将采取什么行动。比如，受害者家属会猜测绑架者是释放还是杀掉受害者，纵然受害者家属最初预测绑架者不会实施绑架。只是因为一个参与者不像你预期的那样行动并不代表你不会对未来发生什么做出判断。

在一个纳什均衡中，一个策略有两个作用：决定规则和推测（这种情况下对策略进行详细规定是很重要的），该策略必须根据参与者的信息集合准确判断出他的行为。一个纳什均衡策略不仅规定了作为一个参与者的决策规则，而且描述了作为另一个参与者对那个参与者决策规则的猜测。

本章总结与练习题

本章总结

一个理性参与者基于对其他参与者行为的信念，选择最大化自身利益的一个策略。这样的一个最佳策略被称为对其他参与者所选策略的一个**最优反应**。如果进一步假设这些推测是准确的（每个参与者正确地预测出其他参与者的策略选择），那么这就是一个**纳什均衡**。纳什均衡的优点是发现了一个适用于所有参与者的“共同满足点”。每个参与者都选择针对其他参与者所选策略的一个最佳策略。

在许多博弈中，重复剔除严格劣势策略（IDSDS）准则缺乏吸引力，因为几乎没有一个策略是严格劣势策略。纳什均衡具有更特别的标准，所以某些博弈中可能只有几个纳什均衡，而有更多的策略组合留存于重复剔除严格劣势策略之中。一般来说，纳什均衡是一个更有价值的博弈求解概念（正是因为上述原因）。然而，我们通过举例分析也发现，一个博弈中可以有许多纳什均衡，一个纳什均衡，或不存在纳什均衡。

求某个博弈中的多重纳什均衡的解时，我们既可以采用直观的方法，也可以运用法则。因而，我们介绍了一个多重纳什求解的常用方法——**最优反应法**（尽管在参与者有多个备选策略的情况下求解效率比较低）。对参与者动机的直觉判断也有助于缩小可能成为纳什均衡的集合范围。

博弈可以分为**纯冲突**类型和参与者有一个**共同利益**的类型。**常和博弈**属于纯冲突类型，因为某一策略使一个参与者受益的同时会使另一个参与者受害。一个典型的例子就是儿童游戏“石头、剪刀、布”，这个游戏同时也是一个**猜谜博弈**，游戏的每个参与者都试图采取其他参与者预想不到的行为方式。处于另一极端的博弈是参与者的利益完全重叠，使一个参与者受益的策略也同样会使另一个参与者获益。行驶规范博弈就表现出了这样的特性，而且该博弈也属于**协调博弈**（当中的参与者只想选择相同的行动）。此外，还有囊括了冲突和共同利益的博弈类型，像打电话博弈、斗鸡博弈和美国偶像迷博弈。在这些博弈中，了解每个参与者的动机（即一个参与者如何对另一个参与者的行为做出最优反应）为推断哪些策略组合会成为纳什均衡提供了一条新途径。

练习题

1.《指环王》（*The Lord of the Rings*）的一个重要时刻是在瑞文戴尔（Rivendell）召开大会，决定由谁把魔戒送到莫多。吉姆利（Gimli）不同意派一个精灵去完成这样的使命，而精灵莱古拉斯（Legolas）也不同意派吉姆利前往莫多。人类博罗米尔（Boromir）反对让他们两个中的任何一个保管魔戒，但他也不是合适人选，因为数千年前正是他的祖先没有抓住机会销毁魔戒，而将它保留了下来。小矮人弗拉多（Frodo）是他们当中最不愿保管魔戒的一个，可他知道必须有人把魔戒扔进莫多火山。将这一剧情改编成一个博弈模型，假设有四个参与者：博罗米尔，弗拉多，吉姆利和莱古拉斯。（当然，还有更多的人物，像阿拉贡 [Aragorn] 等，我们只是把它简化了一下。）每个人对于该由谁承担起护送魔戒的重任有不同的倾向，如表 4—5 所示。

表 4—5　《魔戒三部曲》中各人物偏好排序

人物	第一人选	第二人选	第三人选	第四人选	第五人选
博罗米尔	博罗米尔	弗拉多	没有人选	莱古拉斯	吉姆利
吉姆利	吉姆利	弗拉多	没有人选	博罗米尔	莱古拉斯
莱古拉斯	莱古拉斯	弗拉多	没有人选	吉姆利	博罗米尔
弗拉多	莱古拉斯	吉姆利	博罗米尔	弗拉多	没有人选

除了小矮人弗拉多，每个人都愿意自己承担起这项重任。除了他们自己和弗拉多之外，每个人都不想让其他人拿走魔戒。对于弗拉多而言，他不想承担这项任务，但如果没有人愿意承担的话，他会愿意承担。该博弈中所有的参与者必须同时在四个人选中做出选择。除非大家一致同意（采用全体一致投票原则）某人当选，否则没有人可以完成这项伟大的任务。求出所有的对称纳什均衡。

2. 图4—21所示的是修改行驶规范博弈后的一个博弈模型，当中每个参与者有第三个策略：曲折行进。假设一个参与者选择曲折行进，发生交通事故的概率和另一个参与者选择靠左行驶，或靠右行驶，或也选择曲折行进的概率相同。设定该策略的收益为0，则该收益居于发生冲撞时的收益-1和不发生冲撞时的收益1之间。写出所有的纳什均衡。

	靠左行驶	靠右行驶	曲折行进
靠左行驶	1,1	–1,–1	0,0
靠右行驶	–1,–1	1,1	0,0
曲折行进	0,0	0,0	0,0

图4—21　改编后的行驶规范博弈模型

3. 回到第3章讲的团队合作博弈中，假设一个社团男生和一个社团女生合作，相关的收益矩阵如图4—22所示。写出所有的纳什均衡。

社团女生

		较少努力	适度努力	非常努力
社团男生	较少努力	0,0	2,1	6,2
	适度努力	1,2	4,4	5,3
	非常努力	2,6	3,5	3,4

图4—22　团队合作博弈

4. 思考如图4—23所示的两人博弈。

a. 对于每个参与者，解出重复剔除严格劣势策略之后所剩下的策略。

b. 求出所有的纳什均衡策略组合。

参与者2

		x	y	z
参与者1	a	4,0	2,1	3,2
	b	2,2	3,4	0,1
	c	2,3	1,2	0,3

图4—23

5. 思考如图4—24所表述的两人博弈。

a. 求出重复剔除严格劣势策略之后所剩下的策略。

b. 解出所有的纳什均衡策略组合。

6. 回到第3章中的白人迁徙博弈。假设八座房子中有四座房子由一个房主唐纳德•特朗普拥有，另外四座房子由第二个房主约翰•雅各布•阿斯特拥有。一个策略是承租房子的黑人家庭数量。绘出收益矩阵并推出纳什均衡集合。

参与者2

		x	y	z
参与者1	a	1,2	1,2	0,3
	b	4,0	1,3	0,2
	c	3,1	2,1	1,2
	d	0,2	0,1	2,4

图4—24

7. 回到绑架博弈中，图4—25所示的是该博弈的策略式。写出所有的纳什均衡。

薇薇卡（受害者家属）

		交赎金	不交赎金
盖伊（绑架者）	不绑架/撕票	3,5	3,5
	不绑架/释放	3,5	3,5
	绑架/撕票	4,1	2,2
	绑架/释放	5,3	1,4

图4—25　绑架博弈

8. 伊丽莎白女王决定拍卖王冠宝石，有两个投标人：来自文莱的B先生和来自阿布扎比市的N先生。拍卖规则如下：B先生和N先生同时递交一份书面出价。为了显示自己的与众不同，女王规定N先生的出价必须是1至9（也就是说，必须是1,3,5,7或9）之间的一个奇数（单位是千万英镑），B先生的出价必须是2到10之间的偶数。递交出价最高的投标人获胜，并支付相应的投标价格（如果你能回想起第3章讲的内容，就知道这是一个首价拍卖博弈）。获胜的投标人得到的收益等于估价减去他所支付的价格，而失败的投标人得到的收益为0。假定B先生的估价是8，B先生的出价是7。

a. 以矩阵的形式，写出该博弈的策略式。

b. 求出所有的纳什均衡。

9. 解出图4—26所示三人博弈中的所有纳什均衡。

10. 回到4.4小节中的晋升与阴谋破坏博弈。

a. 推断下面的策略组合是否是一个纳什均衡。（i）参与者1诋毁参与者2，（ii）参与者2诋毁参与者3，（iii）参与者3诋毁参与者1。

b. 求出一个参与者2肯定晋级（概率为1）的纳什均衡。

11. 当存在多个纳什均衡时，一个有效的选择办法是剔除所有涉及一个或多个参与者使用弱劣策略的纳什均衡。图4—17所示的是一个投票博弈，求出该博弈中所有不涉及参与者使用弱劣策略的纳什均衡。

12. 回想一下第 2 章所讲的伽利略和宗教法庭博弈的例子。图 4—27 表述的是该博弈的策略式，写出所有的纳什均衡。

13. 求出图 4—28 所示三人博弈中的所有纳什均衡。

参与者3：A

参与者2

		x	y	z
参与者1	a	1,1,0	2,0,0	2,0,0
	b	3,2,1	1,2,3	0,1,2
	c	2,0,0	0,2,3	3,1,1

参与者3：B

参与者2

		x	y	z
参与者1	a	2,0,0	0,0,1	2,1,2
	b	1,2,0	1,2,1	1,2,1
	c	0,1,2	2,2,1	2,1,0

参与者3：C

参与者2

		x	y	z
参与者1	a	2,0,0	0,1,2	0,1,2
	b	0,1,1	1,2,1	0,1,2
	c	3,1,2	0,1,2	1,1,2

图4—26

教皇乌尔班八世：审阅

审判员

		实施酷刑	不实施酷刑
伽利略	C/C	3,4,5	3,4,5
	C/DNC	3,4,5	3,4,5
	DNC/C	1,5,4	4,2,2
	DNC/DNC	2,1,1	4,2,2

图4—27

教皇乌尔班八世：审阅

审判员

		实施酷刑	不实施酷刑
伽利略	C/C	5,3,3	5,3,3
	C/DNC	5,3,3	5,3,3
	DNC/C	5,3,3	5,3,3
	DNC/DNC	5,3,3	5,3,3

图4—27（续）

参与者3：A

参与者2

		x	y	z
参与者1	a	2,0,4	1,1,1	1,2,3
	b	3,2,3	0,1,0	2,1,0
	c	1,0,2	0,0,3	3,1,1

参与者3：B

参与者2

		x	y	z
参与者1	a	2,0,3	4,1,2	1,1,2
	b	1,3,2	2,2,2	0,4,3
	c	0,0,0	3,0,3	2,1,0

图4—28

附　录

纳什均衡的形式化定义

设想一个博弈有 n 个参与者：1,2,⋯，n。S_i 表示参与者 i 的策略集合；$s'_i \in S_i$ 的含义是策略 s'_i 是策略集 S_i 中的要素。S_{-i} 是 n-1 个参与者（不是参与者 i）的策略（n-1）元组的集合。$V_i(s'_i, s'_{-i})$ 是指当参与者 1 选择策略 s'_i 而其他参与者选择策略 $s'_{-i}=(s'_1,\cdots,s'_{i-1},s'_{i+1},\cdots,s'_n)$ 时他所得到的收益。如果策略 s^*_1 对于其他参与者所选策略（$s^*_1,\cdots,s^*_{i-1},s_i,s^*_{i+1},\cdots,s^*_n$）最大化参与者 i 的收益。那么，对于所有的 i=1,2,⋯,n，一个策略组合（$s^*_1,\cdots,s^*_n$）是一个纳什均衡，更为正式的定义是，当且仅当对于所有的 i=1,2,⋯,n 时，$V_i(s^*_1,\cdots,s^*_n) \geq V_i(s^*_1,\cdots,s^*_{i-1},s_i,s^*_{i+1},\cdots,s^*_n)$ 对所有的 $s_i \in S_i$ 都成立，则（$s^*_1,\cdots,s^*_n$）是一个纳什均衡。

第 5 章

稳定博弈：N 人非连续博弈中的纳什均衡

学习目标

5.1 对称博弈

5.2 非对称博弈

5.3 多重纳什均衡的解

本章导读

在这一章中，我们将研究比两人和三人博弈更复杂的博弈情形。也就是说，研究的范围不仅包括三人以上博弈，也包含参与者人数未知的博弈。后者是指博弈中有 n 个参与者，这个 n 可以是 2,3,10,50,…任何一个你可以想到的数字。无论是几，n 是固定的并且为参与者所知晓。在一些博弈中，不管 n 如何取值，纳什均衡都是相同的，而在其他博弈中，参与者的人数会彻底改变他们的行为选择。

在我们开启新的策略推理海洋航行之前，了解一下我们的行驶航向是很有益处的。在本章的博弈中，所有的参与者都有相同的策略集合，因而也面临着相同的选择。在 5.1 小节中，参与者都有相似的收益函数，因此该博弈是对称的。在接下来的 5.2 小节中，参与者的不同收益函数表明他们的偏好不同。无论是对称的还是非对称的，一些博弈中的均衡行为是指所有的参与者都采取相同的行动。而在其他博弈（甚至是对称的博弈）中，参与者会采用不同的做法。实际上，一个对称博弈不一定有一个对称均衡（如所有的参与者都选择相同的策略）。同样，一个非对称博弈也不一定有一个非对称均衡。

两个重要因素——倾向效应（tipping）**和挤出效应**（congestion）可以决定参与者面对同类型的决策时是否会做出相同的选择。倾向效应是指更多的参与者选择同一策略时，一个参与者也被吸引而选择该策略的倾向。例如，假设你是一个选择穿衣力求大方得体的青少年（这对于你们当中的一些人来说并不难）。在比较 Abercrombie & Fitch 品牌和 The Gap 品牌的衣服时，你的朋友越是选择前者，这一选择对你的吸引力就越大。当某一博弈中出现相当多的参与者选择某一特定策略的情况时，平衡就会被打破，使得所有（或多数）参与者都趋于选择同一策略（因而称之为倾向效应）。近年来，受青年时尚界中存在倾向效应的影响，The Gap 作为龙头服装连锁品牌的地位正逐渐被其他零售商如 Abercrombie & Fitch 所撼动。参与者

PDG

的收益具有这种倾向特质时，极端均衡（当中大部分的参与者都选择相同的策略）的情况就经常发生。谚语“多多益善”就说明了倾向效应的作用。

与倾向效应相反，挤出效应是指越多的人使用同一策略，该策略的吸引力就越小。每个工作日，通勤者在决定上班路线时，都进行着具有挤出效应的博弈。选择某一路线的人越多，该路线的交通也就越拥堵，对通勤者的吸引力就越小。约吉·贝拉（Yogi Berra）在被问及一个很受欢迎的餐馆时，说了一句很经典的话：“没有人再去光顾了，那里实在是太拥挤了。”话语中的讽刺意味正突出了在具有挤出效应的博弈中，均衡所呈现出的特点：并不是每个人都作出相同的选择，因为如果他们这么做的话，就会“太拥挤”。对称博弈的情况更是如此。

在我们分析各种博弈情形时，应当牢记倾向效应和挤出效应。本章中我们的分析内容如下：在操作系统博弈和国内动乱博弈中进行倾向效应分析；见习申请博弈，市场进入博弈和史尼奇博弈（博弈理论课程的必读内容）则用于分析挤出效应。尽管并不是每个博弈都具有倾向效应或挤出效应，但辨别出何时出现这两种效应中的任何一种都有助于确定一个纳什均衡，从而帮助我们快速求出博弈的解。几百万年的进化过程促使人类能够用机智推理代替笨拙的穷举搜索。

5.1 对称博弈

当（1）所有参与者都有相同的策略集合；（2）参与者使用相同策略时的收益相同；（3）交换两个参与者的策略，则他们的收益也被交换时，这个博弈是对称的。例如，在图 5—1 中，参与者 1 和参与者 2 从策略组合（中，低）中得到的收益分别是 2 和 3。对调一下他们的策略使策略组合变为（低，中），则参与者 1 和参与 2 的收益也分别被对调成 3 和 2。这是因为，**对于一个对称博弈中的任何一个策略组合而言，交换策略意味着交换收益。**

		参与者2		
		较低	中等	较高
参与者1	较低	1,1	3,2	1,2
	中等	2,3	2,2	2,1
	较高	2,1	1,2	3,3

图 5—1　对称博弈

改变参与者 2 的策略标签后的情形如图 5—2 所示。显然，改变后的博弈仍然是对称的。我们要认识到参与者 2 选择策略左等于参与者 1 选择策略低。关键的一点是参与者有相同数量的策略，我们可以将他们的策略进行组合，从而满足在上一段提到的收益的两个条件。

在我们分析各种博弈时，假设在一个对称 n 人博弈中，推导出一个非对称纳什均衡（不是所有的参与者都使用相同的策略）是非常有助于求出博弈的解的。如果能推导出一个非对称纳什均衡，那就还有 n-1 个非对称纳什均衡。例如，策略组合（中，低）是图 5—1 所示博弈中的一个纳什均衡，策略组合（低，中）也是一个纳什均衡（因而是对称的）。参与者 1 所选策略中对参与者 2 所选策略低是一个最佳策略的情况类似于参与者 2 所选策略中对参与者 1 所选策略低是一个最佳策略的情形。参与者 2 所选策略低对参与者 1 所选策略中是一个最佳策略的情形与参与者 1 所选策略低对参与者 2 所选策略中是一个最佳策略的情况也相同。因而，若策略组合（低，中）是一个纳什均衡，则策略组合（中，低）也是一个纳什均衡。

		参与者2		
		左	中	右
参与者1	较低	1,1	3,2	1,2
	中等	2,3	2,2	2,1
	较高	2,1	1,2	3,3

图 5—2　图 5—1 中参与者 2 的策略更名后的对称博弈情形



点拨

在对称博弈中，如果一个策略组合是一个纳什均衡，那么参与者对调策略后的策略组合也是一个纳什均衡。也就是说，在对称两人博弈中，若策略组合（s',s''）是一个纳什均衡，则策略组合（s'',s'）同样如此。在对称三人博弈中，若策略组合（s',s'',s'''）是一个纳什均衡，则策略组合（s',s''',s''），（s'',s',s'''），（s'',s''',s'），（s''',s',s''）和（s''',s'',s'）也是纳什均衡。

了解了这一点后，我们能在发现对称博弈中的一个非对称纳什均衡后，推导出还存在其他的非对称博弈，甚至可以确定哪些是非对称博弈。看上去是一个很不错的技巧，是吧？

◯案例研究

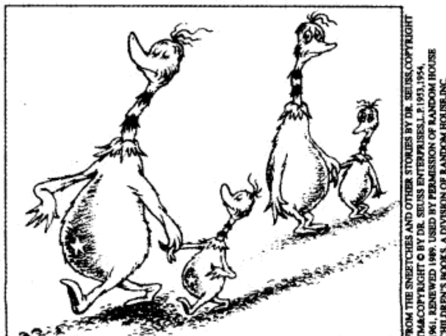
《史尼奇》

故事是这样的，星肚史尼奇的肚皮上有颗星，光肚史尼奇的肚皮上没有星。

在苏斯博士（Dr. Seuss）写的故事《史尼奇》（*The Sneetches*）中，一些虚构人物的肚皮上有一颗星，而另一些的肚皮上则没有星。星肚史尼奇可以享受到吃法兰克福烤香肠的优待，而光肚史尼奇却过着悲惨的生活。或许有些不合理，但史尼奇的世界是稳定的，即星肚史尼奇和光肚史尼奇都不试图改变现状。然而，当斯尔维斯特·麦克猴·麦克豆（Sylvester McMonkey McBean）带着他那架可以给肚皮打上星和去掉星的机器驱车来到这个镇上时，之前的平衡被打破了。你能想起出现了什么样的状况吗？斯尔维斯特·麦克猴·麦克豆的机器又是如何改变史尼奇之间的平衡的呢？

要想探究这个苏斯世界，我们先假设有 n 个史尼奇，且 n 是奇数。此外，起初有不到一半的史尼奇肚皮上有星。随着斯尔维斯特·麦克猴·麦克豆带着他那架可以打上星和去掉星的机器的到来，每个史尼奇必须决定是给她的肚皮上打上星还是去掉星。[①]每个星肚史尼奇必须决定是保留肚皮上的星还是去掉星。类似的，光肚史尼奇必须决定是给肚皮上打上星还是保持光肚皮。

这个故事并没有说明为什么星肚史尼会受到尊敬（确实，故事中决定地位高低因素的选择是很随意的），但是我们假设越稀少的就越珍贵。用 m 表示那天结束时星肚史尼奇的数量。如果 $m<n/2$，那么星肚史尼奇就受到尊敬（因为星肚史尼奇的数量少于光肚史尼奇的数量），而此时，星肚史尼奇的收益是 1，光肚史尼奇的收益是 0。但假如 $n/2<m$，则光肚史尼奇受到尊敬，因而星肚史尼奇得到的收益是 0，光肚史尼奇得到的收益是 1。和故事不同的是，我们假定斯尔维斯特·麦克猴·麦克豆免费提供服务。[②]



星胜史尼奇展示肚皮上的星

注意史尼奇的收益不仅取决于她肚皮上是否有星，也取决于星肚史尼奇的总数。既然打星和去星是免费的，则一个史尼奇是否天生肚皮上有星就无关紧要。表 5—1 表示的是相关的收益。

表 5—1　一个史尼奇的收益

史尼奇肚皮上是否有星	星肚皮史尼奇的总数（m）	收益
有星	$m<\frac{n}{2}$	1
有星	$\frac{n}{2}<m$	0
没星	$m<\frac{n}{2}$	0
没星	$\frac{n}{2}<m$	1

这个群体的稳定格局是什么？是所有的史尼奇都满足于肚皮上有星吗？问题的答案就是推导出的纳什均衡。设想每个史尼奇的策略组合包括打星和不打星，我们的目标是推导出一个策略组合，当中每个史尼奇采取不同的策略不会得到更高收益。

①这里用她这个代词可能有些不可靠，因为我甚至不知道史尼奇是否有性别之分。如果有的话，又是什么。也许史尼奇是雌雄同体或无性的，亦或像两栖动物那样会在一生当中变换自己的性别。也许我应该就此打住（不再坚持符合政治标准），继续我们的故事分析。——作者注

②该故事中，麦克豆给光肚史尼奇肚皮上打星的收费是3美元，给原来的星肚史尼奇肚皮去星的收费是10美元。我猜测之后的收费可能还会上涨。在本章的练习中，你会碰到假设该博弈中麦克豆收费时求出纳什均衡的问题。——作者注

相应的，考虑一个 $m<n/2$ 的策略组合，这时不到一半的史尼奇会选择打星。每个选择打星的史尼奇采取的显然是最佳策略，因为其由此得到的收益是 1（最高收益）。那 $m-n$ 个选择不打星的史尼奇的收益又如何呢？答案是每个人的收益为 0。要判断他们是否可以得到更高收益，需要考虑以下两种情况。

第一种情况是星肚史尼奇的数量不仅小于 $n/2$，也小于（$n/2$）-1，即 $m<$（$n/2$）-1 或 $m+1<n/2$。如图 5—3 所示下面一行 $n=7$ 和 $m=2$ 的情形。如果一个史尼奇原本不打算打星，选择打星后，会使星肚史尼奇的数量由 m 增加至 $m+1$。图 5—3 中上面一行表示的内容就属于这种情况。因为我们已经假设 $m+1<n/2$，所以仍然是星肚史尼奇的数量较少（如图 5—3 所示）。这样一来，一个光肚史尼奇可以从多数光肚史尼奇的一员（得到的收益为 0）转变成为少数星肚史尼奇队伍中的一员（得到的收益是 1）。因为这时任何一个光肚史尼奇都会选择打星，所以这种情况不是一个纳什均衡。总而言之，那天结束之后，星肚史尼奇的数量少于（$n/2$）-1，因而这样策略组合不是纳什均衡。

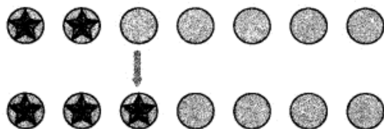


图 5—3　一个光肚史尼奇得到一颗星（$n=7$，$m=2$）

在第二种情形中，我们继续假设 $m<n/2$，这样星肚史尼奇的数量就更少，但现在假定他们的数量超过了（$n/2$）-1，即（$n/2$）$-1<m$。既然我们假设（$n/2$）$-1<m$ 和 $m<n/2$，那么将这两个不等式合并就可以得出不等式（$n/2$）$-1<m<n/2$。但这只表明假设表示星肚史尼奇数量的 m 等于（$n-1$）/2（记得当中的 n 是奇数）。换句话说，m 是小于 $n/2$ 的最大整数。图 5—4 上面一行中 $n=7$，因此 $m=3$ 的情况就说明了这一点。那么，这种情形是一个纳什均衡吗？我们已经知道星肚史尼奇不会得到更高收益。一个光肚史尼奇的收益为 0，因为超过一半的史尼奇没有选择打星。但是因为 $m=$（$n-1$）/2，所以如果一个光肚史尼奇选择打星，则星肚史尼奇的数量变为 [($n-1$)/2]+1 或（$n+1$）/2。有一个史尼奇改变主意，星肚史尼奇的数量就将大于光肚史尼奇的数量，即星肚史尼奇的收益变成了 0。从图 5—4 中下面一行的情况可以看出这一点。选择给肚皮打星并不会增加自己的收益，所以史尼奇满足于选择不打星。

也就是说，当有（$n-1$）/2 个星肚史尼奇和（$n+1$）/2 个光肚史尼奇时，一个光肚史尼奇选择打星和不打星的境况都一样糟糕。一个光肚史尼奇的最佳策略就是仍然光着肚皮。

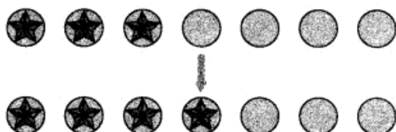


图 5—4　一个光肚史尼奇得到一颗星（$n=7$，$m=3$）

综上所述，星肚史尼奇的数量为（$n-1$）/2 时，是一个纳什均衡。同理，光肚史尼奇的数量为（$n+1$）/2 时，也是一个纳什均衡（自己思考一下原因）。因此，我们推出史尼奇群体的两个稳定格局：或者有（$n-1$）/2 个星肚史尼奇（这时星肚史尼奇受到尊敬），或者有（$n+1$）/2 个光肚史尼奇（这时光肚史尼奇受到尊敬）。

5.1 理解测验

假设一个史尼奇的肚皮上可以刺两种纹身：一颗星或一个横条。总共有7个史尼奇，他们同时决定是纹一颗星，一个横条，还是不纹身。现在史尼奇领地有三个小派别：星肚史尼奇，横条肚史尼奇和光肚史尼奇。设定当一个史尼奇所在派别的数量小于其他派别的数量时，她可以得到的收益是1，否则收益为0。例如，如果有两个星肚史尼奇，两个光肚史尼奇和三个横条肚史尼奇，则每个星肚史尼奇或光肚史尼奇的收益是1。请推断出所有的纳什均衡。

航空安全

安全取决于它最薄弱的环节。如果恐怖分子登上一家国外航空公司的飞机或者盯上一家国外航空公司，这很可能又是一个可怕的悲剧。因此，我们需要确保以有效的方式全面提高安全系数。

——肯尼思·奎因（Kenneth Quinn）
泛美航空 103 号班机安保负责人

因为运送的行李也会接受其他航空公司的检查，所以航空安全不仅取决于一家航空公司的安保措施，也依赖于其他航空公司的安保措施。1988 年泛美航空 103 号班机在苏格兰洛克比（Lockerbie）上空发生爆炸之前，那个装有炸弹的手提箱已经在马耳他接受过安检，然后被转移到了法兰克福和伦敦，之后才被安放在泛美航空 103 号班机上。然而，法兰克福机场和伦敦希思罗机场却没对这个手提箱实施安检。

正如案例开始时所说的那样，航空安全面临的一个最主要的挑战就是它的薄弱环节。如果只有一家航空公司的安全措施不严密，那么这家航空公司就决定着整个航线的安全。一个被称为薄弱环节协调的博弈就捕捉到了这种情形下航空公司的一些动机。下面就研究一些这一博弈的情形。

假设有 $n \geqslant 2$ 家航空公司，每家公司的策略集合是 $\{1,2,3,4,5,6,7\}$，该集合中的每个数字代表每个公司的安保投入，数字越大表示投入的数额越多。令 s_i 表示航空公司 i 所选的策略，并假设与安保措施相关的成本为 $10 \times s_i$。当然，安保措施越严密成本越高。

虽然安保措施的成本由实施该措施的航空公司支付，但利益可以为所有的航空公司共享。假定整个安全系数由薄弱环节决定，即可以由 $50+20 \times \min\{s_1,\cdots,s_n\}$ 决定，其中 $\min\{s_1,\cdots,s_n\}$ 是指 n 个数中的最小值。航空公司 i 的收益是这一公共利益减去个人成本，即：

$$50+20 \times \min\{s_1,\cdots,s_n\}-10 \times s_i$$

在推断纳什均衡之前，先了解一下航空公司的动机。假定 $s_i>\min\{s_1,\cdots,s_n\}$，这样航空公司 i 没有最低安全保障支出。若航空公司 i 减少一个单位的支出（从 s_i 降至 s_i-1），则整个安全系数受到影响，因为 $\min\{s_1,\cdots,s_n\}$ 是不变的（航空公司 i 起初没有薄弱环节）。然而，航空公司的成本减少了 10 单位，从 $10 \times s_i$ 到 $10 \times (s_i-1)$，意味着它的收益增加了 10 个单位。所以，一个航空公司没有最低安保支出时，总是可以通过减少一个单位的措施来增加它的收益。这样做可以在不影响有效安全系数的前提下降低成本。从这一原则可以推断出一个航空公司不采用最低安保支出而选择较高安保措施不是一个最佳策略。因而，在一个纳什均衡中，所有的航空公司必须采用最低安保支出，也就是所有的公司的安保支出必须相同。如果存在纳什均衡，则一定是对称的（注意我们还没有证明存在对称均衡，因为没有非对称均衡）。

如此，我们将集中讨论对称策略组合。假设每个航空公司选择相同的安保措施，用 s' 来表示。因为所有的航空公司面临的决策问题相同，所以我们只需要关注航空公司 1。航空公司 1 所选策略 s' 对其他所有航空公司所选策略 s' 产生的收益是：

$$50+20 \times \min\{s',\cdots,s'\}-10 \times s'=50+20s'-10s'=50+10s'$$

问题是航空公司 1 是否可以通过采取比 s' 更严格或更松的安保措施来获得更高收益。

假设航空公司 1 会选择更严密的安全措施，用 s''（$>s'$）表示。那么所有航空公司的最低安全系数仍是 s'，航空公司 1 的收益是：

$$50+20 \times \min\{s'',s',\cdots,s'\}-10 \times s''=50+20s'-10s''$$
$$=50+10s'-10(s''-s')$$

这一结果必定低于 $50+10s'$。这个公式只是对我们之前所讲内容的一个印证。一个航空公司不想比其他航空公司支付更多的安保费用，是因为这么做不仅不会提高安全系数，反而会增加成本。

下面，假设航空公司 1 选择较弱的安保措施，用 s^0 表示且 $s^0<s'$。最低安全系数从 s' 降至 s^0，那么航空公司 1 的收益是：

$$50+20 \times \min\{s^0,s',\cdots,s'\}-10 \times s^0=50+20s^0-10s^0=50+10s^0$$

这一收益小于 $50+10s'$（因为 $s^0<s'$），即航空公司 1 选择策略 s' 的收益。因而，一个航空公司没有让安保措施低于其他航空公司普通安全系数的动机。

我们推断所有航空公司选择策略 s' 是一个纳什均衡。一个航空公司采用比策略 s' 更严密的安保措施只能增加成本而不会提高安全系数，而采取比策略 s' 更弱的安保措施将使得实际的安全系数低于 s'，抵消掉节省下的安保成本。因为 s' 是一个航空公司策略组合中的一个策略，所以有七个纳什均衡，对应于七个可选策略中的每个策略。

推导出的这一结果既有趣又令人不安，因为纳什均衡意味着所有的航空公司要么采取最严格的安保措施，要么采用最松散的安保措施。而且航空公司并不是对这些纳什均衡持无所谓的态度。一个使每个航空公司增加安保支出的纳什均衡会使所有的航空公司受益。所有航空公司选择安保措施 s' 带给每个航空公司的收益是 $50+10s'$，并且这一收益会随着 s' 的增加而增加。若 $s'=1$，则每个航空公司的收益是 60；如果 $s'=7$，则收益增加至 120。

经济学家通过对大学生进行实验来研究人们在这一情形中的实际行为表现。收益（以美分为单位）与之前的表述相同。比如，所有的学生选择策略 5 的话，则每个人得到的收益是 1 美元。一个人选择策略 3，其他所有的人选择策略 7，则选择策略 7 的人得到 40 美分，选择策略 3 的得到 80 美分。该博弈在一定数量的学生之间进行了 10 轮，所以每个学生都有机会得到 12 美元。有 14 至 16 个人参加了该实验，所以 n 是 14,15 或 16。

合计这七次实验（每次实验的学生主体不同），表 5—2 记录的是第一轮和最后一轮（第十轮）中的选择分布。可以想象得到第一轮有各种各样的选择。一个学生认为其他人很少会选择策略 6，所以他选择策略 6；而另一个不那么自信的学生认为其他人也许会选择几乎不被选择的策略 4，所以她选择策略 4。

表 5—2　薄弱环节协调博弈的实验结果

策略	第一轮中选择该策略的学生比例	第十轮中选择该策略的学生比例
7	31%	7%
6	9%	0%
5	32%	1%
4	16%	2%
3	5%	2%
2	5%	16%
1	2%	72%

在第一轮当中，有 31% 的学生选择了数值最高的策略 7，88% 的学生选择了策略 4 或者更高的策略。只有 2% 的学生选择了数值最低的策略 1。通过对比的方式，让学生作出随机选择，每个策略被选择的概率是 14.3%（有七个策略，$\frac{1}{7}=0.143$）。数值最高的策略 7 被选择的概率高于随机选择的情况。这一结果很令人振奋。但是随着博弈一轮一轮的进行，结果演变成了“底线竞争”。结果越来越倾向于数值较低的策略，这样到第四轮结束时，每一轮数值最低的策略变成了策略 1。到第十轮时，有 72% 的学生选择了策略 1。他们的行为在最差均衡上趋于汇合。[①]

5.2 理解测验

假设有效安全系数由航空公司选择的最高（不是最低）安保措施决定。$\max\{s_1,\cdots,s_n\}$ 表示航空公司的最高策略，我们解出航空公司 i 的收益是 $50+20\times\max\{s_1,\cdots,s_n\}-10\times s_i$。假定策略组合相同，求出所有的纳什均衡。

操作系统：苹果还是微软

> 嗯，你拥有的消费者越多，你的产品价值就越大的理念在当今这个更容易联络和建立交际圈的社会变得更具说服力。我认为这一点也适用于微软的视窗操作系统。随着越来越多的人使用视窗操作系统，如何更好地掌握这个系统似乎成为人们的必修课之一。人们运用微软办公软件的同时，也会使用其附加功能。一个人用微软办公软件制作出一个文档后，想以附件的形式和另一个人进行信息交换的想法在某种程度上是很令人振奋的。正是商业模式中的这些网络效应创造了微软。
>
> ——比尔·盖茨

从 20 世纪 80 年代初开始，发达国家的多数消费者必须在基于微软视窗操作系统的个人电脑与苹果的电脑产品之间进行选择。个人电脑最初是由 IBM 公司研发的，当中配有英特尔公司制造的微处理器和微软公司设计的操作系统（开始是磁盘操作系统，后来推出了视窗操作系统）。相比之下，苹果公司能同时提供硬件和软件组件，起先推出的是苹果二代（Apple II），而后是 Mac。虽然苹果电脑投放市场的时间要早于 IBM 公司，但最终装有微软操作系统的个人电脑占据了市场的主导地位。现今，大约有 95% 的个人电脑安装有微软的 Windows 操作系统。

尽管多数的产品市场和操作系统的市场不同，并不为单一的产品所主导，但当出现网络效应（network effects）时，占优支配的现象就很常见了。如果越多的消费者使用同一产品，该产品对消费者的价值就越大，那么这种产品就具有网络效应。通讯网络中的电话和电子邮件等都是非常典型的例子。以电子邮件为例，如果没有人使用它的话，那它就没有什么价值，它的价值会随着

①在你取消预订的机票之前，不要忘记这些模型和现实之间最大的不同就是不允许参与者进行相互交谈。在博弈前允许相互沟通会使参与者更有可能在最佳均衡上相互协调，即所有的参与者都选择策略7。——作者注

使用人数的增多而增大。操作系统虽然并不直接表现出网络效应，但也受到很大影响。越多的人使用同一操作系统，为该系统编写更多的应用程序（像文字处理器，电子制表软件，游戏等）就越多，从而增加了消费者使用该操作系统的价值。

购买哪种操作系统的决定不是一两个人，十个人或一百个人的博弈，而是几百万人进行的一个博弈。不过，该选择具有策略式选择所有必备的要素，即表现为某人认为最好的操作系统取决于她对其他消费者选择的信念。没有一个消费者会选择使用一个几乎没有什么应用程序的操作系统。

假设有 $n \geqslant 2$ 个人思考购买哪个操作系统。每个人的策略组合由 Mac 和 Windows 组成。假定购买 Mac 操作系统的收益是 $100+10\times m$，其中 m 代表选择 Mac 操作系统的人数。选择 Windows 的收益是 $10\times w$，当中 w 表示的是购买 Windows 操作系统的人数。为了便于分析，设定每个人只能选择其中一种操作系统，那么 $w=n-m$。

需要注意的是一个操作系统的价值会随着购买人数的增多而增大。由产品消费人数，即 Mac 操作系统的消费人数为 $10\times w$，而 Windows 操作系统的消费人数为 $10\times(n-m)$ 决定的一个消费者的收益是衡量网络效应的一个标准。第二个特性是假定 Mac 操作系统更优越的性能得到的收益是 100。

一个纳什均衡会使消费者购买性能更优越的 Mac 操作系统，该操作系统也因而成为了标准。要证明这一点，假设一个策略组合中每个消费者都购买 Mac 操作系统。如果其他 $n-1$ 个消费者购买 Mac 操作系统，那么一个消费者也购买该产品得到的收益是 $100+10\times n$，而购买 Windows 操作系统得到的收益是 10（$=10\times 1$）。因为 $100+10\times n>10$，所以选择购买 Mac 操作系统就是一个最佳策略。由于该博弈和策略组合都是对称的，因而这个推理适用于每一个消费者，所以每个消费者选择购买 Mac 操作系统是一个纳什均衡。

不过，多数人选择购买性能较差的 Windows 操作系统也可以是一个纳什均衡。倘若所有其他消费者都购买 Windows 操作系统，则一个消费者同样选择该系统时得到的收益是 $10\times n$，而购买 Mac 操作系统获得的收益是 110（因为这种情况下 $m=1$）。因而，如果：

$$10\times w \geqslant 100 \text{ 或 } n \geqslant 11$$

那么更好的策略就是和大家一样选择购买 Windows 操作系统。至少有 11 个消费者时，Windows 操作系统从更大网络效应中获得的优势就超过了性能更优越的 Mac 操作系统获得的附加值。因此，每个消费者就会选择购买 Windows 操作系统。

总之，所有消费者选择 Mac 操作系统是一个纳什均衡，但如果 $n \geqslant 11$，那么所有消费者选择 Windows 操作系统的情形也是一个纳什均衡。解出的纳什均衡的结果之所以如此极端（所有的消费者要么都选择 Mac 操作系统，要么都选择 Windows 操作系统）就是因为网络效应本身是倾向效应的一种表现形式。为了更清楚地了解这一点，假定博弈中有 20 个人（即 $n=20$）。图 5—5 表示的是一个操作系统的收益和使用人数之间的关系。横坐标是采用 Windows 操作系统的人数。w 的数值相当于有 w 个消费者选择 Windows 操作系统和 $20-w$ 个消费者选择 Mac 操作系统。纵坐标表示的是一个使用 Windows 操作系统和一个使用 Mac 操作系统的消费者的联合收益。例如，若 $w=14$，则使用 Windows 操作系统的收益是 140（$=14\times 10$），使用 Mac 操作系统的收益是 160（$=100+6\times 10$）。Windows 操作系统使用人数的增多，将导致选择 Windows 操作系统的收益增加而选择 Mac 操作系统的收益降低，从而体现了网络效应的存在。

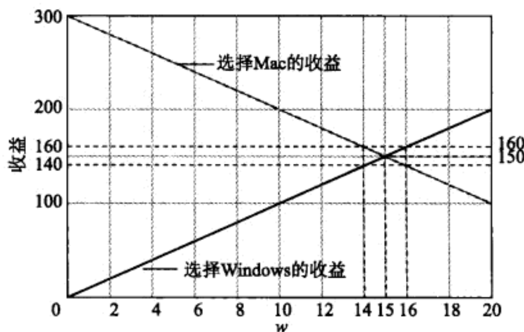


图 5—5　对比 Mac 与 Windows 操作系统

现在假定其他 19 个消费者当中有 15 个选择 Windows 操作系统。那么，一个消费者也选择 Windows 操作系统得到的收益是 160（因为包括她在内共有 16 个消费者使用 Windows 操作系统），而选择 Mac 操作系统获得的收益是 150。这样，她会选择 Windows 操作系统。如果 19 个消费者中有 16 个选择 Windows 操作系统，则选择 Windows 操作系统的倾向性更加明显。这时，选择 Windows 操作系统的收益是 170，而选择 Mac 操作系统

的收益是 140。从图 5—5 可以看出，随着 Windows 操作系统的使用人数从 16 增加至 17 或者更高，Windows 操作系统和 Mac 操作系统之间的收益落差加大。

只要一个消费者认为至少有 15 个消费者选择 Windows 操作系统，那么她也会选择该系统。如果所有的消费者都这么推理，那么事实上，所有的 20 个消费者都会购买 Windows 操作系统，从而确立该系统的主导地位。然而，假如一个消费者认为只有 14 个消费者会购买 Windows 操作系统，则他会选择 Mac 操作系统。图 5—5 清楚地表明，在有 14 个 Windows 操作系统使用者的条件下，选择 Mac 操作系统的收益是 160，而选择 Windows 操作系统的收益是 150（加上这个消费者共有 15 个人选择 Windows 操作系统）。因为预期有更少的消费者会购买 Windows 操作系统，所以 Mac 操作系统的吸引力相对增大。如图 5—5 所示，Mac 和 Windows 操作系统之间的差距会随着 w 数值的减少而增大。

这样，发生倾向效应的临界点就是 15。如果每个消费者都认为有 15 个或更多的消费者会购买 Windows 操作系统，则每个人都会购买这个操作系统。但假如每个消费者都认为有不到 14 个消费者会购买 Windows 操作系统，则每个人都会购买 Mac 操作系统。根据这些信念，市场可以从一个标准转向另一个标准。

一个产品具有网络效应时，消费者的预期是非常关键的。如果多数消费者认为多数消费者会购买某一特定产品，那么他们的信念是会自我实现的，因为所有的消费者都确实会购买那个产品。这样的话，一个公司如何使消费者相信它的产品会受欢迎呢？尽管广告可以实现这一目的，但还并不足以完全实现。因为被某一广告打动的一个消费者都必须知道许多其他的消费者也被同一广告打动。也就是说，他想让这个产品具有竞争力的信息成为共同知识。垃圾邮件（无论是纸质的还是电子的）无法做到这一点。该邮件可能会在许多人中传开，但接收这一邮件的人不知道有多少人同时接收到了这样的邮件。

也许在美国能够实现这一共同知识的最佳途径是在《超级碗》(*Super Bowl*)①电视转播中播放商业广告。之所以这样做不仅仅是因为《超级碗》拥有最高的收视率，更重要的是几乎每个都知道几乎每个人都在观看这一比赛。在 1986 年的《超级碗》中，Discover 信用卡的广告被频频播出，成为了一个具有网络效应的产品的经典例子。只有许多的消费者使用 Discover 信用卡，零售商才会接受它；只有零售商接受了这一信用卡，才会有许多的消费者使用它。因而，有越多的消费者使用这一信用卡，接受它的零售商就更多，从而会吸引更多的消费者去使用这一产品，进而有更多的零售商会接纳这一产品……在 1984 年的《超级碗》中，凭借一个 60 秒钟的商业广告，Mac 操作系统成功地提升了自己的知名度。

实习申请博弈

春季学期过去一半时，你考虑申请一个夏季学期的实习机会。你将可行的实习机会限定在了两个投资银行：纽约的摩根大通公司（JPM）和巴尔的摩的美盛集团（LM）。你和其他人一样都想去摩根大通公司实习，但是你因为下面的两个原因迟迟没有申请：（1）摩根大通公司只给你所在学校的学生提供一个夏季实习的名额，而美盛集团提供的实习名额有 3 个。（2）你知道每个人都很想去摩根大通公司实习，这就意味着要成功获得实习机会的难度会更大。假定由于时间的限制，你只能申请其中的一个实习机会。你会选择哪一个呢？

假设你们学校有 10 个学生对去摩根大通公司或美盛集团实习的机会感兴趣。每个人有相同的偏好，并且每个人都给摩根大通公司的实习机会赋值 200，给美盛集团的实习机会赋值 100。每个人只有在肯定得到摩根大通公司的实习机会时才可以得到 200 的收益，这种情况意味着只有她一个人申请时才可以实现。有一人以上申请该公司的实习机会则会使她的收益减少。通常，申请的人数越多，收益就越少。同样，每个申请美盛集团的人只有在确实得到实习机会时才会有 100 的收益，即在只有 3 个人申请的情况下会得到 100 的收益（该公司只有 3 个实习名额）。申请人数超过 3 人时，申请的收益会随着人数的增多而减少。

表 5—3 详细列出了实习申请博弈的收益情况，图 5—6 是该博弈收益情况的相关绘图。申请摩根大通公司的学生越多，则每个摩根大通公司的申请人的收益越少。摩根大通公司的申请人数越多意味着美盛集团的申请人数越少，每个美盛集团申请人的收益会随着摩根大通申请人数的增多而增大。但在图 5—5 中，越多的消费者

①美国橄榄球联合会和国家橄榄球联合会之间的冠军赛——译者注

选择 Windows 操作系统，每个人的收益就越高。因而，虽然在操作系统博弈中表现出的是倾向效应，在实习申请博弈中体现的却是挤出效应。

表 5—3　实习博弈收益

摩根大通实习人数	每个摩根大通公司申请人的收益	每个美盛集团申请人的收益
0	–	30
1	200	35
2	100	40
3	65	45
4	50	50
5	40	60
6	35	75
7	30	100
8	25	100
9	20	100
10	15	–

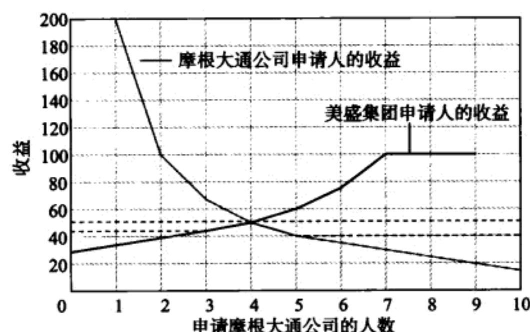


图 5—6　申请摩根大通公司和美盛集团实习机会的收益

在推导出一个纳什均衡之前，先假设没有人申请摩根大通公司。那么，10 个美盛集团申请人中每个人的收益是 30，该收益远远少于一个摩根大通公司申请人所得的收益 200。所以，所有学生申请美盛集团的实现机会不是一个纳什均衡。下面，分析一个策略组合，一个学生申请摩根大通公司，其他 9 个学生申请美盛集团。任何一个美盛集团申请人都可以通过选择申请摩根大通公司来获得更高的收益，因为申请摩根大通公司可以使收益从 35（只有一个摩根大通公司申请人时，每个美盛集团申请人的收益）增加至 100（有两个摩根大通公司申请人时，每个摩根大通申请人的收益）。

让我们拓宽一下分析的范围，假定一个有 m 个学生申请摩根大通公司，一个学生打算申请美盛集团时的策略组合的收益，与有 $10-m$ 个学生申请美盛集团，$m+1$ 个学生中的一个申请摩根大通公司的策略组合的收益。如图 5—6 所示，当 $m<4$ 时，申请摩根大通公司的收益更高，这种情况下这个申请人选择去美盛集团就不是一个最佳策略。只要摩根大通公司申请人数少于 4 人，我们就无法求出一个纳什均衡。

现在考虑另一个极端情况，即假设所有 10 个学生都申请摩根大通公司。那么，每个申请人的收益是 15，这一收益明显小于申请美盛集团所能得到的收益 100。的确，只要有 4 个以上的学生申请摩根大通公司，每个摩根大通公司申请人就可以选择去美盛集团来增加自己的收益。如果这个策略组合中有 m 个学生申请摩根大通公司，那么当 $m>4$ 时，每个摩根大通公司申请人的收益小于 $10-m+1$ 个美盛集团申请人中每个人的收益。这就意味着每个摩根大通申请人应当转换自己的申请。

我们已经证明任何一个人数少于 4 人或多于 4 人申请摩根大通公司实习机会的情形不是一个纳什均衡。这样就剩下了一种可能性，即 4 个学生申请摩根大通公司，6 个学生申请美盛集团。假定每个摩根大通公司申请人和每个美盛集团申请人的收益都是 50。如果一个申请摩根大通公司的学生转而申请美盛集团，那么她的收益会降到 45（见图 5—6）；如果一个申请美盛集团的申请人转而申请摩根大通公司，则他的收益会降到 40。于是，任何一个学生改变她的策略都会使收益减少，因而 4 个学生申请摩根大通公司的情形是一个纳什均衡。

在上述情形中，虽然所有的学生都有相同的选择项和偏好，但他们在这一均衡状态中作出了不同的选择。4 个学生申请摩根大通公司（竞争 1 个实习名额），6 个学生申请美盛集团（竞争 3 个实习名额）。一个对称博弈中出现不对称的行为就是因为出现挤出效应。越多的学生申请一个职位，这个职位的竞争就越激烈（较低的收益就表现了这一点），因而该申请也就不那么有吸引力。

5.2 非对称博弈

参与者有不同的角色，因而就有不同的策略组合，这种情形下的博弈就是非对称的。在绑架博弈情形中（参见第 2 章），绑架者的选择是要么释放受害人，要么杀死受害人；而受害人家属必须决定是否交赎金。即使所有的参与者有着相同的策略组合，但参与者不同的收益情况也会使得这一博弈是非对称的。例如，在选择操作系统时，所有的消费者都面临着相同的选择项，但是他

们会对选择项做出不同的评估。一些消费者十分看重Mac操作系统，而其他消费者则认为不管是什么操作系统，只要是主流的就行。这就是我们下面要探究的非对称的根源，即参与者面对相同的选择项和信息，却拥有不同的收益。我们先看看不同的偏好会产生什么样的结果。

案例研究

市场进入博弈

20世纪90年代末，互联网的兴盛使得当时的网上零售业呈现出飞速发展的势头。越来越多的消费者选择网上购物，与此同时，新的网络服务也不断涌现。图5—7表述的就是这些动态机制的一个代表——B2B（即充当企业对企业进行商品和服务交换媒介的网站）交易数量的增长情况。起初，交易的数量呈现出日益增长的趋势。然而，从2001年初开始，B2B的交易数量急剧下滑。这一模式对于处于发展初期的一个产业来说是非常常见的。图5—8说明的是汽车制造商在20世纪初出现的类似的发展模式，即先是数量猛增而后是下降。

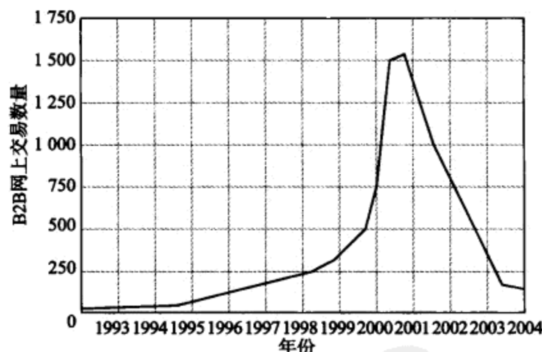


图5—7　1993—2004年B2B网上交易情况

虽然用来解释为什么企业的数量会先上升后下降需要的模型过于复杂，无法在本书中体现出来，但我们至少可以探究进入市场的抉择并预测需要有多少销售商才能够解决企业的进入和退出争端。

假设刚刚出现了一个新的市场，比如20世纪90年代中期出现的网上图书市场（亚马逊网上书店从1995年7月开始进行网上图书销售）。假定有五个公司会进入这个市场，分别是公司1,2,3,4和5。要在这个市场运作，每个公司必须支付进入成本（你可以把该成本看成是创建一个网站的成本），整合组织结构和进行网站宣传。

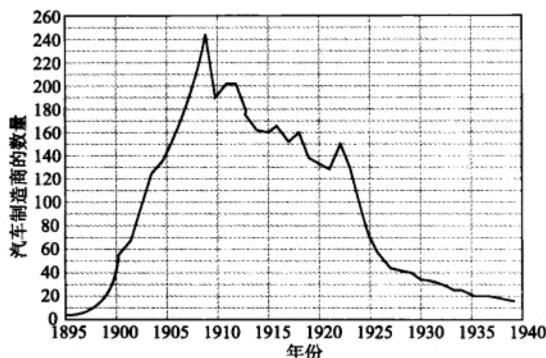


图5—8　1895—1938年汽车制造商的情况

表5—4呈现的是未来市场进入者所面对的不同进入成本。公司1是最高效的，因为它的进入成本只有100，①而公司5是效率最低的，进入成本高达210。一个公司可能会凭借负责人自己编程或者是作为一个传统零售商的子公司（如bn.com进入市场时就是Barnes & Noble的一个子公司），使自身有机会借鉴现有技术的优势，实现较低的进入成本。

表5—4　不同的市场进入成本

公司	进入成本
1	100
2	160
3	180
4	200
5	210

在描述过市场进入成本后，思考一下进入市场的收益情况。每个零售商预期进入市场后会获利。然而，有更多的竞争网站时，每个公司的利润就会降低。假定所有的零售商获得的总利润相同（剔除进入成本之前的毛利润）。表5—5记录的是每个公司会获得的总利润以及该利润和零售商数量之间的关系。如果只有一个零售商进入这一市场，则这家公司就有垄断权，可获得1 000利润。相反，有两家公司进入的话（就出现了双头买主垄断），每家公司的利润是400。需要注意的是第二个零

①不用管单位是什么，因为它们可能是几千美元，几百万谢克尔或几十亿列克。——作者注

售商的出现使总的工业利润由1 000降到了800。出现这种情况不仅是因为两个零售商会分割市场，而且争夺消费者的竞争会促使他们降低价格，从而减少了利润。如果有三个零售商（产生三头垄断），那么每家公司的利润是250，依此类推。

表5—5　所有进入公司的总利润

进入公司数量	每家公司的利润	总的工业利润
1	1000	1000
2	400	800
3	250	750
4	150	600
5	100	500

因为这些公司同时决定是否进入市场，那么每个公司的策略组合就由进入和不进入构成。一个公司不进入的收益为零，而进入的话收益是它的净利润（总利润减去进入成本）。比如，如果公司2和其他两家公司选择进入，则公司2的收益是90（=250-160）。基于表5—4和表5—5提供的信息，表5—6汇总了各种收益情况。

表5—6　一家公司进入市场的收益

公司	其他进入市场的公司数量				
	0	1	2	3	4
1	900	300	150	50	0
2	840	240	90	−10	−60
3	820	220	70	−30	−80
4	800	200	50	−50	−100
5	790	190	40	−60	−110

请注意，没有一个策略是严格劣势策略，尽管公司1选择进入的策略弱优于选择不进入的策略。当有三家或更少公司进入市场时，选择进入策略产生一个严格更高收益，这一收益同样适用于有四家公司进入市场的情况。对于公司1之外的每一家公司而言，选择进入是应对两家或更少公司进入市场时的一个最优反应，但选择不进入是应对三家或四家公司进入市场时的一个最优反应。因而，每家公司的任何一个策略都是应对其他四家公司某一策略组合的一个最优反应。

下面要求出一组纳什均衡。与之前对某一特定策略组合的判断不同，我们先分析在一个纳什均衡中有多少进入者。

■存在一个没有进入者的纳什均衡吗？不存在，因为第一个进入者的总利润是1 000，这超出了所有公司的进入成本。因而，如果所有其他公司选择不进入，一个公司选择进入就是一个最优策略。因此，所有公司选择不进入的策略组合不是一个纳什均衡。

■存在一个只有一个进入者的纳什均衡吗？答案还是不存在，因为第二个进入者的总收益是400，这超过了所有公司的进入成本。所以，在一个策略组合中如果一个公司进入而其他四个公司不进入，那么这四家公司中的任何一家都可以通过进入市场而增加自己的收益。因而只有一家公司选择进入的策略组合不是纳什均衡。

■存在一个只有两个进入者的纳什均衡吗？按照此前的论证方式（我们已经很熟悉了），可以得出不存在的答案。假定在一个策略组合中有两家公司选择进入，其他三家选择不进入。那么每家选择不进入的公司收益是零。现在仔细观察表5—6中的竖列2。所有的收益都是正数，所以无论哪三家公司选择不进入，一家公司选择进入策略都可以得到正数收益。

■存在一个只有三个进入者的纳什均衡吗？确实存在。事实上，存在六个这样的纳什均衡：（1）公司1、2和3选择进入；（2）公司1、2和4选择进入；（3）公司1、2和5选择进入；（4）公司1、3和4选择进入；（5）公司1、3和5选择进入；（6）公司1、4和5选择进入。以第一个策略组合为例。每个选择进入的公司都可以得到一个正数收益，即公司1的收益是150，公司2的收益是90，公司3的收益是70。选择进入策略就是一个最佳策略。对于选择不进入的公司4和5来说（如果你细心观察表5.6中的竖列3会发现），选择进入策略会产生负收益：即公司4的收益是-40，公司5的收益是-50。这样，公司1、2和3选择进入策略而公司4和5选择不进入策略的策略组合是一个纳什均衡。我们可以用同样的推理方法证明其他五个策略组合是纳什均衡，以及证明三家公

司选择进入（比如公司 3、4 和 5）的任何一个策略组合不是一个纳什均衡。

我是如何证明这六个策略组合是纳什均衡的呢？不是通过穷举搜索，而是通过思考这些公司的动机。首先，考虑如果三家公司选择进入，而当中不包括公司 1 的话，那么这样的策略组合就不可能是一个纳什均衡，若公司 1 不是这三家进入公司，则它的收益为零，但选择进入的话，得到的收益就会变成 50。因而，如果有三家公司选择进入，那么公司 1 必定会出现在其中。其次，注意观察有两家公司选择进入时（竖列 2 中所有的收益都是正数）其他四家公司中的任何一家都可以从选择进入获得一个正收益；而有三家公司选择进入时（竖列 3 中的所有收益除公司 1 的以外都是负值），他们中的每一家也选择进入得到的是一个负收益。这样，我们就可以推断出公司 1 和其他两家公司选择进入的任何一个策略组合都是纳什均衡。

存在一个有三家以上公司选择进入的纳什均衡吗？有四家进入者时，每家公司的总利润是 150，但只有一家公司的进入成本不超过 150。所以，如果四家公司选择进入，他们当中至少有一家的收益为负值，意味着这家公司最好不要选择进入。这一推理也可以从表 5—6 得到印证，即竖列 3 和 4 中只有一家公司的收益不是负值。这样，有三家以上的公司选择进入的策略组合就不是纳什均衡。

纳什均衡的一个有趣的特性是效率最高的公司不一定是选择进入的公司。最高效的纳什均衡是公司 1，2 和 3 选择进入，因为他们的进入成本最低。可是，也存在公司 4 和 5 而不是 2 和 3 选择进入的策略组合是纳什均衡的情况。假定预期公司 4 和 5 选择进入（和公司 1 一起），选择进入策略对公司 2 和 3 来说不会产生利润，即使与公司 4 和 5 调换位置之后，公司 2 和 3 也会比公司 4 和 5 产生更多的利润。

5.3 理解测验

把市场进入博弈中的公司1剔除，只剩下公司2,3,4和5同时决定是否选择进入市场。相关的收益情况仍和图5.6所表述的内容相同。写出所有的纳什均衡。

国内动乱

一小部分人是如何压制一个大群体的？除了掌控媒体和经济资源以外，军队的支持也非常重要。然而，更重要的一点可能是民众协调失败。在非民主社会中，掌权的总是少数人。如果民众站起来反抗统治机构的话，独裁者是没有生存空间的。问题的关键是如何使民众协调一致，从而实现民众示威。

假设一个国家有 500 个国民，每个人正在考虑是否参加示威活动。示威得到的利益是 $50 \times m$，其中 m 表示参加示威的人数。这一利益设计体现了更大规模的抗议更有影响力的合理理念。示威的利益为所有参与者所共享。

示威的代价是牺牲个人的利益，比如面临监禁的危险。这里，我们假设人们对付出的代价有不同的衡量标准。为了便于分析，简单设定这个国家的民众分为三类（见表 5—7）。当中有 100 名积极分子，他们参加示威的代价最小，即 6 000；进步分子有 100 名，每个人参加示威的代价是 8 000；还有 300 名资产阶级分子，他们想参加反抗，但由于所处的状况并不太糟糕，所以革命一旦失败，遭受的损失最大。对他们而言，参加示威付出的代价是 20 000。

表 5—7　国内动乱博弈

国民分类	国民数量	个人代价	临界示威点
激进分子	100	6 000	120
进步分子	100	8 000	160
资产阶级分子	300	20 000	400

一个人不参加示威的收益是零，而参加示威的收益是利益 $50 \times m$ 减去他付出的代价。换句话说，当且仅当：

$$50 \times m - 6\,000 \geq 0$$

时，一个激进分子才会参加示威。对上述这个表达式求解之后，我们得出 $m \geq 120$。因而，如果一个激进分子认为另外有 120 人会参加示威，那么她也会参加并且得到收益 50（$=50 \times 121 - 6\,000$）。（如果预期另外 119 个人参加示威，那么她参加示威的收益是零并且她也不在乎是否参加示威。）同理，一个进步分子预期 m 满足不等式：

$$50 \times m - 8\,000 \geq 0 \text{ 或其等价式 } m \geq 160$$

时，会参加示威。这样，因为要付出更高代价，所以一个进步分子在预期参加示威的人数大于一个激进分子预

期的参加人数时，才会参加示威。最后，预期有 400 人参加示威时，资产阶级分子才会选择参加示威，即：

$$50 \times m-20\,000 \geqslant 0 \text{ 或其等价式 } m \geqslant 400$$

引起一个人参加示威的最小示威规模称之为她的临界示威点。表 5—7 给出的就是相关信息。

在求出一个纳什均衡时，一个进步分子参加示威是一个最佳策略，那么一个激进分子参加示威也是一个最佳策略的假设有助于我们求出结果。也就是说，示威人群中如果有进步分子，也一定有激进分子。要证明这一假设，先回想当一个进步分子预期有 160 人参加示威时，他认为选择参加示威是一个最佳策略，而一个激进分子在有 120 人参加示威的前提下，就会参加示威。因而，如果有足够多的示威人群吸引进步分子参加示威，那么也能吸引激进分子参加示威。直觉上，一个进步分子和一个激进分子实现了相同的利益，但是由于后者的代价较小，所以一个进步分子示威时，一个激进分子总是会选择示威。按照相同的推理方式，可以判定如果一个资产阶级分子认为选择参加示威是一个最佳策略，那么激进分子和进步分子选择参加示威也是一个最佳策略。

从以上的分析中，我们可以列出有可能成为纳什均衡的策略组合是：（1）没有人参加示威；（2）只有激进分子参加示威；（3）只有激进分子和进步分子参加示威；（4）每个人都参加示威。没有人参加示威的情况显然不是纳什均衡，因为如果一个激进分子预计没有人参加示威，则他参加示威的收益是-5 950（=50×1-6 000），这比待在家里不参加示威的情况还要糟糕。这时，参加示威对一个进步分子和一个资产阶级分子来说就更没有什么吸引力了。

只有激进分子参加示威的情形又如何呢？假如所有的激进分子都参加示威，则总共有 100 名示威者，每个激进分子的收益是-1 000（=50×100－6 000）。这并不能够使参加示威成为一个最优策略，因为在只有激进分子参加示威前提下，他们的人数是不足以支撑一个均衡示威的。可是，如果激进分子和进步分子都参加示威，那么一个进步分子的收益是 2 000（=50×200－8 000）。因而，所有的进步分子都出去参加示威，并且我们知道激进分子也会同他们一起进行示威。相反，一个资产阶级分子则不想参加这样的示威，因为他这么做得到的收益是-9 950（=50×201-20 000）。所以，只有激进分子和进步分子参加示威的情形是纳什均衡。

最后，每个人参加示威的情形也是纳什均衡。如果一个资产阶级分子预期其他所有的人都会参加的话，那么他加入到他们当中来的收益是 5 000（=50×500－20 000），这使得参加示威成为一个最佳策略。因而，激进分子和进步分子参加示威也是最佳策略。

总而言之，纳什均衡可以包括不发生示威活动，有 40% 的民众参与的小规模示威活动以及全民参与的大规模示威活动。一场示威活动深入社会的程度取决于民众的预期。如果他们预期会发生大规模的示威，那么就会发生大规模的示威活动；如果预期没有民众参与示威，那么就不会发生示威活动。发生示威活动的关键因素在于人们的预期和对民众预期会有大规模人群参加示威活动信念的协调。

5.3 多重纳什均衡的解

本章研究的许多博弈情形都存在多重纳什均衡。多重纳什均衡会引发问题，因为均衡点越多，我们对行为的预测就越不准确。这一问题引起了博弈理论界的高度重视，我们也会在本书中继续关注这一问题。在这一节中，我们会介绍如何解决多重博弈过程中普遍存在的问题。①

区分各种类型的多重性是很重要的，因为虽然一些类型的多重性是无关紧要的，但另外一些类型却能产生很大的影响力。思考电话博弈中出现的多重纳什均衡（参见 4.2 小节），图 5—9 重现了这一内容。该博弈中有两个纳什均衡，一个是科琳选择打电话，而温妮选择等电话；另一个是温妮选择打电话，而科琳选择等电话。不论选择哪一个均衡，博弈的结果都是一样的，因为在这两个均衡中都是一个人选择打电话，另一个人选择等电话，尽管每个人的角色会发生改变。我们不需要过多关注这一类型的多重性（还有一个需要注意的事项，我们稍后会提及），因为所有的均衡对结果的预测都大致相同。当中出现类似于这种多重纳什均衡的博弈还包括斗鸡博弈（参见 4.1 小节）和行驶规范博弈（参见 4.2 小节）。

然而，在某一类型的多重纳什均衡中，除选择策略的

①这部分可以看做第5章所讲内容的一个附录。对第5章所列博弈的参考已经用题外话这样的事标识做出明显的区分。——作者注

参与者身份各异之外，策略分配完全相同时，我们需要推断：有一个以上的纳什均衡时，参与者是如何在某一特定的均衡中进行协调的？科琳和温妮是怎样协调谁先打电话的？①斗鸡博弈中司机是如何协调谁先转弯的？如果参与者重复博弈，那么就会产生一些共同的信念组合，从而在某一特定的均衡中出现博弈分歧。但假如我们的任务是在他们第一次博弈时就预测行为的话，这个办法就行不通了。也就是说，这种情形下博弈是无解的。

		温妮	
		打电话	等电话
科琳	打电话	0,0	2,3
	等电话	3,2	1,1

图 5—9　电话博弈

☞ 题外话

这种类型的多重纳什均衡也出现在史尼奇博弈和申请实习博弈中。在后一博弈中，4 个学生选择申请摩根大通公司和 6 个学生选择申请美盛集团的任何一个策略组合都是纳什均衡。具体而言，实际上有 210 个纳什均衡，因为使 10 个学生中有 4 个申请摩根大通公司的分配方式有 210 种。这两种博弈都存在挤出效应，而且多重非对称纳什均衡确实普遍存在于在具有挤出效应的对称博弈中。因为使用一个策略的收益会随着使用该策略人数的增加而减少，这种情形常常出现在一个非对称均衡中。在一个对称博弈中，如果有一个非对称均衡，那么就会有更多的非对称均衡，同时这些非对称均衡都有着相同的策略集合（有多少人使用策略 1，就有多少人使用策略 2 等类似的情况），而他们之间唯一的不同只是表现在哪个参与者会使用哪种策略。

改变一下行驶规范博弈模型，假定其中的两个司机习惯靠右行驶（如图 5—10 所示）。该博弈中有两个纳什均衡，即两人都选择靠右行驶和两人都选择靠左行驶，但基于得到的收益不同，博弈的结果也是不同的。尽管靠左行驶也是一个纳什均衡，但两人会选择靠右行驶。这一类型的多重纳什均衡也出现在美国偶像迷博弈和投票博弈中（参见 4.4 小节），类似的，博弈的结果也大不相同。这种情形下，不同的纳什均衡会做出不同的行为预测。

		司机2	
		靠右行驶	靠左行驶
司机1	靠右行驶	5,5	1,1
	靠左行驶	1,1	2,2

图 5—10　两名美国司机参与的行驶规范博弈

☞ 题外话

这一类型的多重性经常出现在具有倾向效应的博弈中。所有参与者都选择某一特定策略的情形可以成为一个纳什均衡，而他们都选择另一策略的情形也可以成为一个纳什均衡，因为一个策略的吸引力在很大程度上取决于它的大众性②，操作系统博弈的情形就是如此。每个人选择 Mac 操作系统的策略组合是一个均衡，每个人选择 Windows 操作系统也是一个均衡，然而，消费者会选择后一个均衡。

多重纳什均衡做出不同的行为预测时，博弈理论可以提供几个技巧，从而帮助我们从存在于一些博弈中的纳什均衡集合中求出唯一的纳什均衡。思考图 5—11 所示的博弈情形。这个博弈中有三个纳什均衡：一个是斯库特和布朗都选择策略低；另一个是斯库特选择策略高，布朗选择策略中；第三个是斯库特选择策略中，布朗选择策略高。在你试图选择策略组合（低，低）之前（因为这个策略组合是对称的，并且该博弈也是对称的），需要注意的是策略低弱劣于策略高。因而，一个谨慎的参与者是不会选择策略低的。

		布朗		
		低	中	高
斯库特	低	2,2	1,1	0,2
	中	1,1	3,3	1,4
	高	2,0	4,1	0,0

图 5—11

一个选择技巧是优先考虑**非劣纳什均衡**（undominat-

①经济学家乔尔·索贝尔（Joel Sobel）建议说，在现实生活中，应该是先打电话的人回电话，因为这个人肯定知道另外一个人的电话号码。——作者注

②即选择该策略人数。——译者注

ed Nash equilibria），即参与者不选择弱劣策略的纳什均衡。这一方法可以剔除策略组合（低，低），这样还剩下两个纳什均衡。回到投票博弈中（参见 4.4 小节），可以推出有五个纳什均衡，但除了一个均衡之外，在所有的均衡中一个或更多的选民使用弱劣策略。在这一博弈中，只有一个非劣纳什均衡，即股东 1 和股东 2 赞同方案 B 与股东 3 赞同方案 C。我建议将证明上述推断作为一个练习。

另一个经常使用的选择准则基于**收益占优**（payoff dominance）和经济学家所称的**帕累托标准**（Pareto criterion）。如果其他策略组合中不存在每个参与者有一个严格更高收益，那么这个策略组合就满足**收益占优**的条件。尽管纳什均衡基于个人理性，即每个参与者会针对其他人的行为作出对自己最有利的行为，但收益占优是基于集体理性，即选择一个最有利于每一个人的策略组合。这一选择准则是将收益占优和纳什均衡相结合。只运用前者意味着选择一个违反个人理性的策略组合。这一选择准则被称为**均衡收益占优**（equilibrium payoff dominance），即关注纳什均衡组合当中没有其他的均衡出现每个参与者有一个严格更高收益的情况。符合这一特性的纳什均衡就叫**收益占优纳什均衡**（payoff-dominant Nash equilibrium）。

按照纳什均衡收益占优准则，我们只需关注图 5—10 所示博弈中的一个均衡（靠右，靠右），因为这一均衡带给两个司机的收益要高于均衡（靠左，靠左）的收益。运用这一准则，我们可以求出美国偶像迷博弈唯一的解是三个女孩都穿印有字母的 T 恤。

题外话

均衡收益占优标准可以在航空安全博弈求解中发挥很大作用。该博弈有七个纳什均衡，并且可以按照收益进行由高到低的排序。在所有的纳什均衡中，只有其中出现所有航空公司选择最严格安保措施的一个纳什均衡可以使每个人都得到最高收益。这一准则也可以用于操作系统博弈，从而求出唯一的解。

使用纳什均衡收益占优准则的一个基本原理是如果参与者事先联系和沟通，那么对一个具有共同利益的纳什均衡进行协调对大家都有好处。他们知道除这一均衡之外，对任何事达成**协议**（agreement）都是不可靠的，也就是说，有一个或更多的参与者会背离达成的这一协议。因而，参与者应该是现实的，只考虑那些**自我实施协议**（self-enforcing agreement），即基于个人利益选择遵从，因而不需要像法庭和警察这样的组织去强迫实现。在所有的自我实施协议中（如纳什均衡），所有的参与者都同意他们应当只关注收益占优的协议。

谜题

图 5—12 所示的博弈情形中有两个纳什均衡：策略组合（上，左）和策略组合（下，右）。其中，策略组合（上，左）是一个收益占优纳什均衡。然而，对参与者 1 而言，策略下弱优于策略上；对参与者 2 来说，策略右弱优于策略左。因而，策略组合（下，右）是一个非劣纳什均衡。你认为哪一个纳什均衡更令人信服呢？

		参与者2 左	参与者2 右
参与者1	上	3,3	1,3
	下	3,1	2,2

图 5—12　收益占优与弱占优

在其 1961 年出版的著作《冲突的战略》（*Strategy of Conflict*）中，诺贝尔奖获得者托马斯·谢林（Thomas Schelling）引入了**焦点**（focal point）概念。同色情和爵士的情形类似，我们很难从形式上明确定义出什么是焦点。就像前任最高法院大法官波特·斯图尔特（Potter Stewart）在一次司法裁决中所说的："我无法向你证明什么是色情，但我看到它就会认出它。"也可以用传奇爵士乐手刘易斯·阿姆斯特朗（Louis Anmstrong）的话解释："如果你必须知道什么是爵士乐，那么你永远也无法了解它。"尽管焦点不容易被定义，但它仍是一个非常有说服力的概念。

要了解什么是焦点以及如何用它来帮助解出一个纳什均衡，我们就必须先确保博弈不是发生在真空中，而是发生在现在丰富的社会情境中。所有的参与者都是一个拥有历史的人，这一历史当中的一部分是他们的共同知识。虽然策略被贴上了诸如"上"、"下"这样的标签，但在实际的情境中，它们都有着具体的内涵，例如，"在

车站碰面”、“棒球场上见”。根据参与者的身份和他们可能会采用的策略，可能会有一个显著的纳什均衡，因而成为一个当然的选择或焦点。但要使所有的参与者都认同这一选择，每个人都认为这一选择是显而易见的还不够，还需要让这个显著性本身成为共同知识。

为了搞清楚这一理念是如何运作的，我们就以驾驶博弈为例来进行分析。假定这一博弈中，关键的问题是选择在哪边的人行道上走。在伦敦，因为法律规定靠左驾驶，所以焦点也许是靠左走。的确，事实似乎也是这样，我在英国的人行道上的许多尴尬经历就证明了这一点。

5.4 理解测验

求出国内动乱博弈中的收益占优纳什均衡。

本章总结与练习题

本章总结

在本章研究的博弈情形中，所有的参与者都面临着相同的选择。不论是一个消费者选择购买哪种操作系统的电脑，一个公民决定是否参加示威，一个史尼奇是否给肚皮上打星，还是一家公司选择是否进入某一市场，都是如此。许多分析都没有明确说明有多少参与者。一些策略组合（如购买 Mac 操作系统）不论有多少参与者都是纳什均衡，而其他策略组合（如购买 Windows 操作系统）是否是纳什均衡取决于参与者的人数。

如果除了所有的参与者拥有相同的选择之外，还假定他们有相同的收益，那么该博弈就是对称的。一个对称博弈不一定是每个人都采取相同行动的均衡博弈。而且能否产生一个对称均衡是由收益函数是否具有倾向效应或是挤出效应决定的。在倾向效应中，更多的参与者使用同一策略时，使用这一策略的相对可能性就更大；在挤出效应中，更多的参与者使用同一策略时，使用这一策略的相对可能性就更小。具有倾向效应的博弈常常会导致出现（几乎）每个人做出相同举动的纳什均衡。例如，如果有足够多的人购买 Mac 操作系统，那么选择也成为一个 Mac 操作系统使用者的倾向性会更大，因为这就意味着为这一操作系统编的软件会更多，你可以很轻松地与其他使用者交换文档。相反，当出现挤出效应时，面对相同选择的人们会做出不同的抉择。比如，申请实习机会的学生并不都想申请同一个投资银行，因为竞争会很激烈。在一个纳什均衡中学生会申请不同公司的实习机会。虽然面对相同的情境，但人们的行为表现会有很大差异。

纳什均衡的多重性是本章所分析的博弈中共同拥有的一个特性。确实，每个博弈中的纳什均衡都不止一个。从更广泛的层面来说，博弈理论经常会碰到多重性。我们提出了一些在多重均衡中如何进行选择的准则，从而求出博弈中唯一的解。其中的一些准则包括**非劣纳什均衡**（这一方法可以剔除当中出现参与者使用弱劣策略情形的纳什均衡集合），**收益占优纳什均衡**（由此可以剔除当中出现每个人得到最高收益的一个纳什均衡所在的均衡集合）。一些博弈可以得出唯一的解，但在许多博弈中却无法实现。

如果一个准则总是可以求出唯一的解，那么发现这一准则无疑就像博弈理论得到了“圣杯”一样。一些博弈理论家认为这么做是不现实的，而且以此作为追求目标只能把自己变成堂吉诃德。然而，其他理论家则认为这一任务是极富挑战性的，完成这一使命的人简直可以和亚瑟王相媲美。不管你的观点是什么，所有的理论家都非常看重这样的准则，因为做出准确的预测和解释才是博弈理论的本质所在。

练习题

1. 第 84 期《科学》杂志计划举行一场比赛，任何人都可以递交一个要求得到 20 美元或 100 美元的申请。如果有不到 20% 的申请要求得到 100 美元，那么每个人都可以得到他们申请的相应数额。如果有超过 20% 的申请要求得到 100 美元，那么每个人都一无所获。尽管因为伦敦的劳埃德集团（Llogda）不愿承担损失，该比赛计划泡汤，但是我们仍然可以分析哪一个纳什均衡会做出准确预测。假设有 10 000 人可能会参加这一比赛，并且收益以金钱计算。

a. 假定每一个潜在的参赛者都递交了 20 美元或 100 美元的申请。求出一个纳什均衡。

b. 假定只有购买第 84 期《科学》杂志才可以得到一个申请表，该杂志的单价是 21.95 美元。那么，每个人的策略组合包括不购买该杂志，购买该杂志并递交一个 20 美元的申请，以及购买该杂志并递交一个 100 美元的申请。假设杂志的价值是零，求出一个纳什均衡。

c. 将杂志单价的假定改为 19.95 美元，其余情形与 b 部分的描述相同。求出一个纳什均衡。

2. 某人在聚会上掏出 100 美元，宣布对其进行拍卖。聚会上有 $n \geqslant 2$ 个潜在的出价人。100 美元的所有者提出如下的拍卖程序：所有的出价人同时递交一个书面出价。每个人（不仅仅是最高出价人）支付自己所报的出价，最高出价的出价人得到这 100 美元（假设最高出价是正数）。如果有 m 人递交最高出价，那么每个人获得这 100 美元的 $\frac{1}{m}$。每个人的策略组合是 $\{0,1,2,\cdots,1\,000\}$，所以出价可以高达 1 000 美元。假如 b_j 表示参与者 j 的出价，$\max\{b_1,\cdots,b_n\}$ 表示 n 个出价中的最高出价，那么参与者 i 的收益是：

$$-b_i,\quad b_i<\{b_1,\cdots,b_n\}$$
$$\frac{100}{m}-b_i,\quad b_i=\{b_1,\cdots,b_n\}$$

其中，m 是出价等于 $\max\{b_1,\cdots,b_n\}$ 的出价人人数。求出所有的纳什均衡。

3. 现在是哥伦比亚地区早上较为拥堵的通勤时间，有 100 个司机，每个司机正在决定是走收费公路还是走乡间小道。收费公路的通行费是 10 美元，而乡间小道不收费。在选择路线时，每个司机只关心收入（用 d 表示）和通行时间（用 t 表示）。如果一个司机的最后收入为 y，通行时间为 t，那么他的收益就被假定为 $y-t$（当中一个单位通行时间的价值等于 1 美元）。当天开始时每个司机的收入是 1 000 美元。如果有 m 个司机选择收费公路，那么每个司机在收费公路上的通行时间就被假定为 m 美元。相反，如果 m 个司机选择乡间小路，那么他们在乡间小路的通信时间就是 $2m$ 美元。司机们同时决定是走收费公路还是乡间小路。

a. 求出每个参与者的收益函数（即根据每个参与者的策略组合说明她收益情况的表达式）。

b. 解出一个纳什均衡。

4. 回到 5.2 小节中所讲的史尼奇博弈。在 n 个史尼奇中，k 表示星肚史尼奇的数量且 $k<\frac{n}{2}$。现假设斯尔维斯特·麦克猴·麦克豆去掉星（如果史尼奇天生肚皮上有一颗星）或打上星（如果史尼奇天生肚皮上没有星）的收费是 p。假定 $0<p<1$。除了如果选择打上星或去掉星时，必须得扣除费用 p 之外，一个史尼奇的收益情况和表 5—1 的表述相同。求出所有的纳什均衡。

5. 假设几个朋友出去吃饭费用均摊。问题是某人可能会点较贵的菜，因为他知道多出的部分费用会由其他人买单。为了分析这一情境，假设有 n 个用餐者，并且食物偏好相同。表 5—8 表述的是菜单所列三种菜中每种菜的价格和每个人对每种菜的估价。估价由一人愿意支付的最高饭钱来衡量。

表 5—8　就餐困境博弈　　单位：美元

菜名	价值	价格	盈余
意大利面	21	14	7
鲑鱼	26	21	5
牛排	29	30	−1

盈余等于支付的饭钱减去菜价。意大利面的价格是 14 美元，每个用餐者对他的估价是 21 美元。因而，如果一个用餐者必须支付所点菜的全部价格，那么每个人都可能会选择意大利面，因为这个菜的盈余大于鲑鱼和牛排的盈余。但事实上，用餐者会径直选择标价 30 元的牛排，如表所示它的盈利为负数。一个参与者的收益等于她所点饭菜的估价减去她必须支付的价格。后者假

定等于账单总额除以就餐人数。例如，如果有三个用餐者，每个人各点了一个不同的菜，那么点意大利面的用餐者得到的收益是：

$$21-\left(\frac{14+21+30}{3}\right)=21-21.67=-0.67$$

点鲑鱼的用餐者得到的收益是：

$$26-\left(\frac{14+21+30}{3}\right)=4.33$$

点牛排的用餐者得到的收益是：

$$29-\left(\frac{14+21+30}{3}\right)=7.33$$

果然，点的菜越贵得到的收益越多，因为所有人支付的费用相同。

a. 假设有两个用餐者（n=2），他们会怎么点菜（在一个纳什均衡中）？

b. 假设有四个用餐者（n=4），他们会怎么点菜（在一个纳什均衡中）？

6. 还是分析5.2小节中所讲的市场进入博弈，但是需要假定五个潜在的市场进入者的市场进入成本相同，即都是300美元。假定进入市场公司的总数不变，表5—9描述的是如果一个公司选择进入获得的净利润（或收益）。和之前的情形一样，不进入市场的收益为零，每个公司可以选择进入还是不进入。写出所有的纳什均衡。

表5—9　公司选择相同策略时的博弈情形

公司数量	每个公司的总利润	每个公司的净利润
1	1 000	700
2	400	10
3	250	−50
4	150	−150
5	100	−200

7. 在一个治安较乱的街区有 $n \geq 2$ 个居民。每个居民必须决定是否从事偷窃的犯罪活动。如果一个人选择成为一名小偷，并且不被警察抓住，那么他得到的收益是 W；如果被警察抓住，得到的收益是 Z。如果他选择不当小偷，得到的收益是零。假定 $W>0>Z$。所有 n 个居民同时决定是否从事偷窃。一个小偷被抓的可能性等于 $\frac{1}{m}$，其中 m 表示选择从事偷窃的居民数。那么，犯罪活动越多，被抓的可能性就越小，警察就要调查更多的犯罪案件。假设其他 $m-1$ 个居民也选择当小偷，那么一个小偷的收益是：

$$\left(\frac{m-1}{m}\right)W+\left(\frac{1}{m}\right)Z$$

求出所有的纳什均衡。

8. 分析如图5—13所示的博弈情形。

a. 写出所有的纳什均衡。

b. 如何求出这些多重纳什均衡中唯一的均衡。

		参与者2		
		x	y	z
参与者1	a	2,2	1,2	0,0
	b	2,1	2,2	1,3
	c	0,0	3,1	0,0

图5—13

9. 分析如图5—14所示的博弈情形。

a. 写出所有的纳什均衡。

b. 如何求出这些多重纳什均衡中唯一的均衡。

		参与者2		
		x	y	z
参与者1	a	1,0	1,2	0,1
	b	0,0	2,1	3,3
	c	1,2	1,1	0,1

图5—14

10. 假设一个国家有 n 个公民，v_i 表示公民 i 预期参加示威的价值。列举每一个公民，使得公民1预期参加示威的价值大于公民2的预期价值，公民2的预期价值大于公民3的预期价值，以此类推，即 $v_1>v_2>\cdots>v_n$（=0），其中公民 n 预期参加示威没有价值。假定每个公民参加示威付出的代价相同，都是 $\frac{c}{m}$，当中 $c>0$，m 是参加示威的人数。那么，公民 i 参加示威得到的收益是 $v_i-\frac{c}{m}$，而不参加示威得到的收益是零。设定 $v_i-c<0$。求出所有的纳什均衡。

11. n 个医学预科学生计划参加医学院入学考试。每个学生必须决定在参加考试之前是否选修一门预备课程。x_i 表示学生 i 的选择，当 $x_i=0$ 时，表明她不会选修该课程，而当 $x_i=1$ 时则表明她会选修该课程。一个学生关心她在医学院入学考试中的成绩和是否选修预备课程。s_i 表示学生 i 的医学院入学考试成绩，r_i 代表她在 n 个参加考试学生中的排名。具体而言，r_i 等于1加成绩严格高于学生 i 成绩的学生人数。为了清楚地说表示这一规则，举三个例子加以说明。若 $s_i \geq s_j$ 对所有的 $j \neq i$ 都成立，则 $r_i=1$。换句话说，如果没有人的成绩高于学生 i 的成绩，那么她就排名第一。若 $s_i<s_j$ 对所有

的 $j \neq i$ 都成立，则 $r_i=n$。也就是说，如果学生 i 的成绩最低，那么她的排名就是 n。若 $s_1>s_2>s_3=s_4>s_5$，则 $r_1=1$，$r_2=2$，$r_3=r_4=3$，$r_5=5$。现在假设学生 i 的收益等于 $b(n-r_i)-x_i c$，当中 $b>c>0$。需要注意的是，一个学生选修预备课程的成本等于 c 且如果一个学生的排名每向前增加 1，那么她的收益会增加 b。学生 i 的成绩假设由公式 $s_i=a_i+x_i z$ 决定，其中 $a_i>0$ 且 $z>0$。a_i 和学生的天赋有关，也是她不选修预备课程会得到的成绩。如果她选修预备课程，那么她的成绩就会增加 z。假定：

$$a_1>a_2>\cdots>a_{n-1}=a_n$$

这就意味着学生 1 比学生 2 聪明，学生 2 比学生 3 聪明……学生 n-2 比学生 n-1 聪明，学生 n-1 和学生 n 一样聪明。最后的假设是：

$$a_i+1+z>a_i, i=1,2,\cdots,n-1$$

在这一博弈中，有 n 个学生同时决定是否选修医学院入学考试的预备课程。求出一个纳什均衡。



第 6 章

稳定博弈：连续博弈中的纳什均衡

学习目标

6.1 不使用微积分求解纳什均衡

6.2 运用微积分求解纳什均衡（选读）

本章导读

在本书已经探究过的博弈中，策略的数量是极其有限的。在某些情境中，出现较少的策略是很正常的。例如，一个司机决定靠左行驶还是靠右行驶（参见第 2 章讨论的行驶规范博弈），或选择转向还是坚持冲下去（参见第 2 章的斗鸡博弈）；一个公民决定是否参加示威活动（第 5 章的国内动乱博弈）；一个公司选择是否进入一个市场（第 5 章）。而在其他一些情境中，有很多的选择项似乎更合乎情理。比如，我们研究拍卖出价（第 2 章）时，将一个出价人的可行出价选择严格限定在五个方案上，而实际上，他可以有许多的出价选择。同样，在团队合作博弈（第 3 章）中，一个学生的努力程度被划分为低、中和高三个等级，但如果我们以他花费的时间作为衡量努力程度的标准，那就会有许多的等级。特别是一个人在选择金钱额度或时间量时，她通常都会许多的备选项。

在这一章当中，我们将分析博弈当中出现许多策略的情形。这个“许多”是什么意思？生活在博茨瓦纳（Botswana）的布什曼人（Bushmen）对数的认识只包括一到六，而超过六的所有数字都被划分到许多的行列。这个对于“许多”的定义几乎无法实现我们的目的。因为我们对“许多”的理解要远远大于七，事实上，它可能是无限大的。当然，由于无限个策略过于复杂而很难精确记录，所以我们假定策略可以从低到高排序，一个策略集合是一组数字。一个数字代表一个价值，这一价值可以是指工作时间，也可以是像一个公司产品的质量这样不太容易计量的事物。只要能够对各选择项进行排序（如按照从低到高的顺序），就可以赋予其相应的数字，无疑，较大的数字表示较高的质量。

无限数集合的一个代表就是自然数集合，如 1,2,3…。可是，我们对“许多”的定义还有一个特性，即两个策略之间存在第三个策略。而自然数并不具有这样的特性，例如 73 和 74 之间就没有自然数。具有这一特性的数集合是实数集合，比如 2 和 10 之间的所有数字（不仅仅是整数），例如 [2,10] 这个区间的数就包括像 5.347, 8/3 和 π（3.141 592 653 589 793 238 462 643 383 279 502 884 197…）。

实数包括有理数和无理数。一个有理数可以用由两个整数组成的一个分数表示。[①]例如，$\frac{1}{3}$是一个有理数，1.721也是有理数且与分数$\frac{1\,712}{1\,000}$相等。一个无理数不能由一个分数表示，它小数点之后有无限多个数，并且不会循环，[②]如 π 就是一个无理数。所有有理数和无理数的集合就是**实数集合**。

在这一章中，一个策略集合或者是一个实数区间或者是整个实数集合。这样的策略集合被称为**连续策略集合**（Continuous strategy set），该集合中没有空缺的数字。你也许在想为什么需要这么多的策略？原因是虽然有理数集合很大，但是使用起来并不方便，因为当中包含了许多间隙。将一个策略集合设定为一个实数区间的最大好处是我们可以运用 17 世纪伟大的数学发现——微积分，这一点会在 6.2 小节中的分析中得到很好的体现，但在 6.1 小节中我们不需要使用微积分。

在 6.1 小节中，我们会研究两个经典的竞争模式，一个是经济学领域中百货商场通过调节价格争夺消费者，另一个是政治学范畴中候选人通过调整竞选纲领参加总统选举。也许你会认为在无限个策略组合中推导出一些纳什均衡要比大海捞针更难，但是通过机智的推理分析还是能够使问题迎刃而解（虽然还是会有一些可能性无法排除）。6.2 小节虽然是选读内容，但也是本章的重点，在这节中我们会用微积分的方法轻而易举地推断出纳什均衡。先是介绍一个常规方法以及运用该方法进行博弈求解必须满足的条件。然后使用这一方法分析市场竞争以及人类行为是如何导致猛犸象灭绝的。最后研究配额资助的原理和如何运用这一原理来增加慈善捐款。

6.1 不使用微积分求解纳什均衡

在有无限个策略的情况下，用穷举搜索的方法求解纳什均衡是行不通的。但思考每一个策略组合是否属于纳什均衡的方法永远奏效，而且我们还可以做得更好。尽管没有一个法则可以在博弈中有无限个策略的条件下求出该博弈的解，但是我们仍可以采取手段使得某些情形中的任务量大大削减。

这一技巧就是理解博弈中参与者是如何决策的。我们要站在参与者的角度去思考他有什么样的动机。假如你就是那个参与者，你会怎么做？参与者是如何增加自己的收益的？参与者是怎样对其他参与者的行动做出最优反应的？一旦对参与者所处情境有了一定的认识，你就会有更好的洞察力，可以剔除很多很多的策略组合。

要具备这样的理解力，一个常用的方法是深入研究问题。选择一个策略组合（任何一个都行），然后分析它是不是一个纳什均衡。除非你十分走运，否则这个策略组合很可能不是纳什均衡。这时你要思考为什么它不是纳什均衡，因为得到的答案也可能适用于其他的策略组合。这一方法看上去似乎很神秘，但当你看到它在一些案例中的运用时，就会对这一方法有更深刻的理解。

○案例研究

相同产品的价格竞争

假设你在尼亚加拉大瀑布景区度假，准备给家人和朋友挑选一些纪念品。你在克利夫顿山庄的纪念品商店转了一圈，发现里面卖的东西大同小异：尼亚加拉瀑布钢笔（转动钢笔时桶里的小人就会跑到瀑布上面），尼亚加拉瀑布冰箱磁铁，尼亚加拉瀑布雪花玻璃球等。当你在每家商店随意观看时，会发现当中商品的标价都相同。是偶然？是幻觉？抑或是竞争？究竟是为什么？

要分析这一情况，不妨假设有两个纪念品商店出售相同的商品。为了便于计算，只讨论其中的一种商品。每个商店以 10 美元的单价从同一个制造商那里订货。例如，如果商店 1 卖出 13 件这样的商品，那么总成本就是 130 美元。商店 2 的总成本也是 130 美元。

顾客查看了发现两个商店的商品相同之后，就会从价格最低的商店那里购买（如果两个商店的定价都太高，顾客可能就不会购买）。假设价格越低，购买商品的顾客就越多。价格和销售数量的关系由市场需求曲线的概念加以概括（如图 6—1 所示）。如果两个商店的最低定价是 p，那么销售的数量就是 $100-p$。

①这里使用的“有理”术语不同于我们之前对行为“理性”的表述。——作者注

②譬如，$\frac{1}{4}$等于0.25，小数点后的数字是有限的；$\frac{1}{7}$等于0.142 857 142 857 142 857……，小数点后的数字重复序列142 857。——作者注

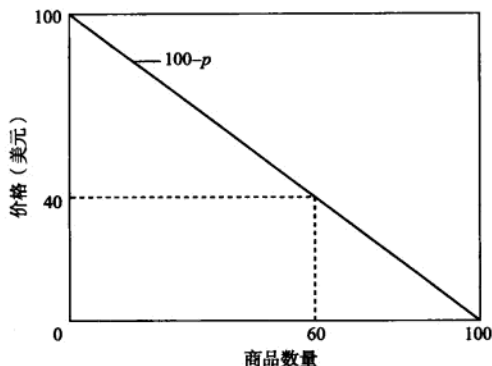


图 6—1　纪念品市场需求曲线

下面定义什么是一个商店的需求。因为顾客只根据价格作出选择，若商店 1 的定价低于商店 2，即 $p_1<p_2$，则所有的顾客只从商店 1 购买商品，商店 1 的销售量是 $100-p_1$。可是若商店 1 的定价高于商店 2，即 $p_1>p_2$，则没有顾客会从它那里购买商品，所以也就没有销售量。最后如果两个商店定价相同，那么总需求会在他们之间平分。比如，两个商品都要价 40 美元，那么顾客想买 60 件（=100−40）商品，每个商店的销售量是 30 件。概括这一表述之后，发现商店 1 的需求曲线用 $D_1(p_1,p_2)$ 表示为：

$$D_1(p_1,p_2)=\begin{cases}100-p_1, & p_1<p_2\\ \left(\frac{1}{2}\right)(100-p_1), & p_1=p_2\\ 0, & p_2>p_1\end{cases}$$

商店 2 的收益曲线定义与之类似。

我们通过绘图来进一步分析商店 2 定价 40 美元时，商店 1 的需求曲线（如图 6—2 所示）。注意曲线图中价格在 40 以上的猛增部分。定价高于 40 时，商店 1 卖不出商品；定价低于 40 时，可以卖出全部的 40 件商品；定价为 40 时，可以得到一半的市场需求，或卖出 30 件商品（用实心点表示，以区别于空心点）。

要确定一个商店的策略集合，不妨设一个策略是商品价格，并且可以在 [1,100] 区间内取值。此外，假设每个商店只在乎赚钱；商店 1 的收益等于取得的销售收入——即商品的单价乘以销售的数量，或 $p_1\times D_1(p_1, p_2)$，减去它的成本——销售数量的 10 倍，或 $10\times D_1(p_1, p_2)$。因而商店 1 的收益（或利润）函数是：

$$\begin{cases}(p_1-10)(100-p_1) & p_1<p_2\\ \left(\frac{1}{2}\right)(p_1-10)(100-p_1) & p_1=p_2\\ 0 & p_2<p_1\end{cases}$$

图 6—3 表示的就是当 $p_2=40$ 时，商店 1 的收益函数。同理，商店 2 的收益函数是：

$$\begin{cases}0 & p_1<p_2\\ \left(\frac{1}{2}\right)(p_2-10)(100-p_2) & p_1=p_2\\ (p_2-10)(100-p_2) & p_2<p_1\end{cases}$$

这一博弈被称为贝特朗价格博弈，是由约瑟夫·贝特朗（Joseph Bertrand）在 1883 年提出的。

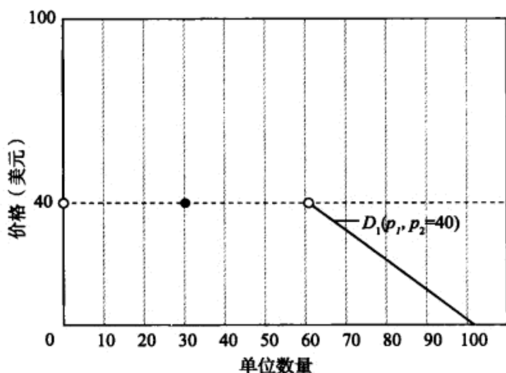


图 6—2　商店 2 定价为 40 美元时商店 1 的需求曲线

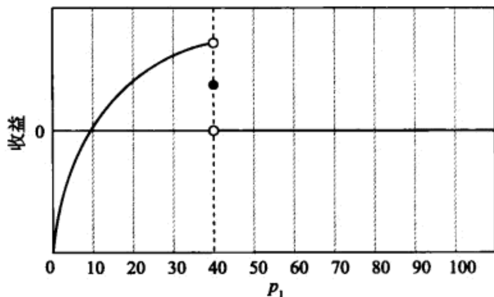


图 6—3　商店 2 定价为 40 时商店 1 的收益函数

一个潜在的纳什均衡可以是处于 [0,100] 区间的任何价格组合。我们现在看能否剔除一些可行的策略集合。如果一个商店对产品的定价小于成本 10 并且销售量不定，那么它的收益就是负值。这样的一个定价不可能是

最佳策略，因为不论另一个商店的定价是多少，一个商店定价 10 的策略得到的收益为零。因而，假如最低的定价不到 10 美元，采用最低定价策略也不是最优反应。由此，我们可以推断任何一个商店定价都低于 10 的策略组合不是纳什均衡。

我们成功地剔除了有可能成为一个纳什均衡的无限个策略组合。现在唯一的问题是还剩下无限个潜在的纳什均衡策略组合，即定价等于或大于 10 的策略组合。为了简化分析，我们把重点放在对称策略组合上，也就是，两个商店定价相同的策略组合。

思考任何一个对称策略都大于 10 的情况。例如，设定两个商店的定价都是 $p'>10$。图 6—4 所示就是另一个商店定价为 p' 时，商店 i 的收益函数。这种情况下，商店 i 以 $p'-10$ 的单价卖出了 $\left(\frac{1}{2}\right)(100-p')$ 件商品，因而得到一个正收益 $\left(\frac{1}{2}\right)(p'-10)(100-p')$。我们将这一定价和定价为 $p'-\varepsilon$（$\varepsilon>0$）的情况做一比较。若 ε 的数值很小，则定价为 p' 和定价为 $p'-\varepsilon$（如比较 $p'-10$ 和 $p'-\varepsilon-10$）时，每销售一件商品获得的收益相同。可是，定价为 $p'-\varepsilon$ 时，销售量几乎是定价为 p' 时销售量的两倍，我们可以比较 $\left(\frac{1}{2}\right)(100-p')$ 和 $100-(p'-\varepsilon)$。因而，当 ε 的数值很小时，定价为 $p'-\varepsilon$ 所得的收益大于定价为 p' 的收益，即：

$$[(p'-\varepsilon)-10][100-(p'-\varepsilon)]>\left(\frac{1}{2}\right)(p'-10)(100-p')$$

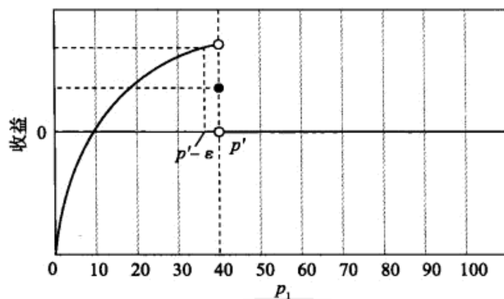


图 6—4　另一商店的定价为 p' 时，商店 i 的收益函数

上述的推论揭示出一个商店以低于对手的价格出售商品的强大动机，即相对于和对手价格持平的情况，这样做不仅几乎不会减少销售每件商品所获的利润，还使得自己的需求量翻倍。但是只要一个商店低价出售，另一个商店就不会满足于现状，因为他将没有销售量。因此，任何一个定价超过成本 10 的对称价格组合都不是纳什均衡，因为两个商店的动机都是以较低的价格销售。

简而言之，我们先推断低于 10 的对称价格组合不是一个纳什均衡，然后推导出高于 10 的对称价格组合不是一个纳什均衡。现在只有一种可能性，即两个商店的定价都是 10。这样的对称价格组合是纳什均衡吗？因为具有对称性，所以我们只需要考虑其中的一个，商店 1 即可。对手的定价为 10 时，商店 1 定价也是 10 的销售量是 45 件，但因为以成本价销售所以收益为零。如果定价高于 10，那么将销售不出去商品，因为所有的顾客都会从商店 2 那里购买，所以商店 1 的收益还是零。如果定价低于 10，虽然可以占领整个市场，但却是在赔本销售，即其收益为负。因而，如果对手以成本价定价，那么商店 1 也只能以成本价销售。至此，我们推导出了一个纳什均衡。

6.1 理解测验

假定在克利夫顿山庄有三个商店销售同类商品。假设三家定价相同时，每家获得三分之一的市场需求，剩余的情形与之前的描述相同。求出所有的对称和非对称纳什均衡。

这一结论给我们的启事是虽然只有两个商店，但竞争已经相当激烈了。以成本 10 定价才能使他们勉强经营。如果定价低于 10，那么其中的一家会选择停业以免遭受损失。这一情形正是消费者梦寐以求的理想状态。可是，对于消费者来说，生活并不总是那么美好。对于这一点，我们会在第 14 章中从更加符合现实的角度去分析百货商店如何应对竞争并维持一个较高的销售价格。

谜题

如果你点击进入迪克体育用品（Dick's Sporting Goods）的一些网址，就会发现这些产品不仅价格相同（竞争在起作用），而且产品销售编号也相同（思考为什么）。这究竟是怎么回事？（这一谜题有解。）

价格一致承诺化解价格竞争

百思买集团："商家郑重承诺——如果您在一家百思买的商场购物时，发现其他商家同

类商品的定价比我们的还要低，请告知我们，我们会马上调整为相同价格。”

环城百货：“价格承诺——如果您看到其他商家同类商品的宣传价格更低，请告知我们。经查属实，我们将以九折的价格给您更大的实惠。”

一些零售商家常用的一个策略是如果消费者发现其他商家对同类产品的定价比他们的更低，那么就会与更低价格持平。这种被称为价格一致承诺的做法已经被电器商店，办公用品超市，超级市场，轮胎经销商和其他许多的零售商所实践。表面看来，这种做法是很富有竞争力的行为，但这样不会压低价格吗？实际上，这么做会抬高价格！了解一点博弈理论后，你就会相信这一点。

设想之前所举例子中的两个商店销售同一类产品。假定这两个商店中的每一家都实施价格一致承诺策略。也就是说，当 $p_1 \leqslant p_2$ 时，商店 1 会以 p_1 的价格销售商品（这样他的标价就不会比对手的高），可是当 $p_1>p_2$ 时，他会以 p_2 的价格销售商品（这样低价承诺开始生效）。虽然商店 1 可能会以 p_1 的价格出售商品，但实际上商品可能会按照两家定价中较低的价格出售，可以用 $\min\{p_1, p_2\}$ 表示 p_1 和 p_2 的最小值。

那么，商店 1 的收益函数是：

$$[\min\{p_1, p_2\}-10]\left(\frac{1}{2}\right)[100-\min\{p_1, p_2\}]$$

类似的，商店 2 的收益函数是：

$$[\min\{p_1, p_2\}-10]\left(\frac{1}{2}\right)[100-\min\{p_1, p_2\}]$$

因为低价承诺的影响，两个商店最终都以相同的价格（一方提出的最低价格）销售商品，每家各占 50% 的销售量。看起来似乎商店就应当以低价销售。但在得出结论之前，我们需要仔细分析一个商店在定价时的动机。

先推断如果街道上只有一个商店时，该商店会怎样定价也许有助于我们之后的分析。在上述情况下，这个商店的收益函数是：

$$(p-10)(100-p)$$

图 6—5 表述的也是相同的收益情况。可以看出这一收益函数呈山形，在价格为 55 时达到了最高值。作为垄断商的这家商店会把价格定在 55，销售 45 件商品，得到的收益是（55−10）（100−55），即 2 025。

下面将重点关注如何求出对称纳什均衡。假定一个策略组合中两个商店的定价都是 p'，并且 $10 \leqslant p' \leqslant 55$，这样价格就处于成本和垄断价格之间。图 6—6 描述的是当 $10<p'<55$ 时，商店 1 的收益函数。虚线表示的是商店 1 是垄断商时他的收益情况，实线代表的是假定商店 2 的定价是 p' 时商店 1 的收益。$p_1 \leqslant p'$（如商店 1 的定价不高于商店 2 的定价）时，商店 1 以 p_1 销售，得到总利润（p_1−10）（100−p_1）的一半，实线位于虚线和零的正中间（商店 2 因为低价承诺的影响也会以 p' 的价格销售）。$p_1>p'$ 时，商店 1 以 p' 的价格销售（因为低价承诺的影响，会和商店 2 的定价持平），得到的收益是 $\left(\frac{1}{2}\right)(p'-10)(100-p')$。这样，收益函数呈水平状，因为无论标价 p_1 是多少，商店 1 最终都会按商店 2 较低的定价 p' 销售商品。

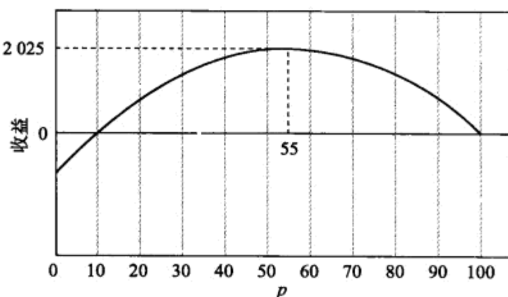
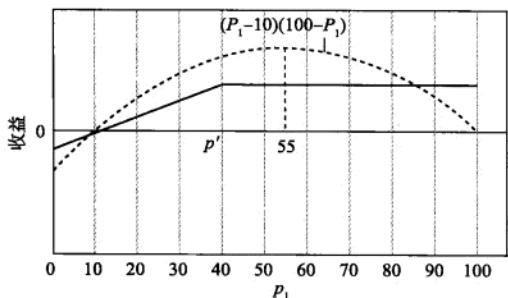


图 6—5　克利夫顿山庄只有一家商店时这家商店的收益函数



虚线表示的是两家商店定价为 p_1 或商店 1 是垄断者时的总收益，实线表示的是商店 2 定价为 p' 并且两家商店价格一致时商店 1 的收益。

图 6—6　克利夫顿山庄有两家商店时每家店的收益函数

从图 6—6 可以看出任何一个等于或高于 p' 的价格都是商店 1 的一个最优反应，因为此时它的收益是最大的。特别是 p' 是一个最优反应。因为策略组合和博弈都是对称的，所以 p' 也是商店 2 的一个最优反应。因而，

不论 p' 是 10,55 还是两者之间的任何一个价格，两个商店以 p' 定价是一个纳什均衡。

这样，商店实施承诺价格一致的策略时，就会有许多的对称纳什均衡。可是，如果我们关注收益占优纳什均衡，就会发现求出的解是两个商店以 55 定价，因为这一策略组合是使两个参与者收益达到最高的纳什均衡。但请注意 55 是一个垄断价格。在承诺价格一致的影响下，两个商店之间的竞争消失，他们最终的定价和只有一个商店时的定价相同。

要弄清为什么会出现这样一个令人惊讶的结果，不妨回到之前不存在承诺价格一致的模式中去，思考商店的动机是什么。对于两个商店而言，高于成本的相同定价不是一个均衡，因为每家都有压低对方定价的动机。削价虽然会略微减少每件商品的利润，但更会使一个商店的销售量翻倍。然而，在承诺价格一致的影响下，削价根本不起作用。即使商店 1 的定价比商店 2 的低，也许已经打算从商店 2 购买的消费者仍然会做出相同的选择，因为商店 2 会和商店 1 较低的价格持平。商店 1 削价的结果只能是以更低的价格把商品出售给同一批顾客。因而，承诺价格一致摧毁了压低对手价格的动机，从而使商店维持较高的价格。看上去是加强竞争的做法反而消解了竞争。

一项对超级市场采用承诺价格一致的研究表明这种做法确实提高了商品价格。在 1983 年，一个北加利福尼亚的食品商 Big Star 采用了价格一致承诺的做法并每周印发传单，其中列出了 9 000 多种商品的价格。对于所列这些商品的价格，Big Star 承诺会和他的主要竞争对手 Food Lion 的价格相同。在 1985 年，另一个竞争对手 Winn-Dixie 也推行类似的做法，承诺和 Food Lion 的价格相同。博弈理论预测传单中所列商品的价格会高于那些没有出现在传单上的商品的价格，因为所列商品的价格（而不是后者）会受到价格一致承诺的影响。尽管商品价格受到的平均影响并不大（2% 左右），但还是能感受到这种影响。比如，在采用承诺价格一致之前，麦式速溶咖啡在 Food Lion 的售价是 2.19 美元，在 Winn-Dixie 的售价是 2.29 美元，在 Big Star 是 2.33 美元。但采用这一做法之后，三家食品商都将麦式速溶咖啡的售价调整为 2.89 美元。

竞选总统

我曾观看最后一轮总统竞选辩论。就在辩论快要结束时，我萌发了换频道的念头，因为无聊的内容着实让我看不下去。乔治·布什和约翰·克里所说的内容似乎都没有什么新意，都是一些早已听过的内容。在辩论接近尾声时，我不禁感到自己像是在观看《爱丽丝漫游奇境》（*Alice & Wonderful*）中在爱丽丝周围跳舞的孪生兄弟叮叮和咚咚，只不过在这种情境之下称他们为小哑巴（Tweedle-Dumb）和大哑巴（Tweedle-Dumber）也许更合适一点。

总统候选人会在许多方面展开竞争，如进行电视宣传、寻求政治资金、发表巡回演说等，但最有效的竞争手段也许是他们看待问题的立场。为了说明这一竞争方式，假设只有一个问题，对问题的立场用 [0,1] 区间的数字表示。比如，这个问题是有关税收的，数字越大表明税收越高。或者这一问题是关于如何在福利事业与国防开支之间分配资金，数字越大意味着国防开支所占比例越大而福利事业所占比例越小。亦或设想候选人并不明确表明立场，而是更多阐释自己的理念。以 0 表示的立场意味着"极左"（即很激进），以 1 表示的立场意味着"极右"（即很保守），而中间立场由分数 $\frac{1}{2}$ 表示。

为了描述该博弈的策略式，假设参与者是候选人 D（民主党）和候选人 R（共和党），每人都有一个策略集合 [0,1]。x_i 表示候选人的立场或策略，$i=D,R$。从颇为世俗的角度假设候选人只在乎是否当选，而不关心自己的立场。要描述出一个候选人的收益函数，需要说明策略（如两个候选人的立场）是如何决定竞选结果的。要做到这一点需要花费一点功夫。

假设每个选民有一个处于立场集合 [0,1] 之间的理想立场，而且许多选民根据他们的理想立场均匀分布在 [0,1] 之间。换句话说，最倾向于立场 0.3 的选民数量和最倾向于立场 0.72 的选民数量相同，以此类推。假定所有选民都投票，并且将选票投给所表明立场和自己的理想立场最接近的候选人。这样，一个候选人所表明立场的可信度（即选民相信他所说的话）是毫无疑问的。也许这一理想假设会抵消掉对候选人偏好的世俗假设。

现在分析候选人D的立场位于候选人R立场的左边，即$x_D<x_R$时，博弈的情形（如图6—7所示）。在候选人D立场左边的选民（即他们的理想立场低于x_D）会给她投票，因为x_D要比x_R更接近于他们的理想立场。同理，在x_R右边的选民会给候选人R投票。在x_D和x_R之间的选民会如何投票呢？假设那些处于中间的不偏不倚的选民所处的位置是$(x_D+x_R)/2$。因为处在$(x_D+x_R)/2$的选民和任何一个候选人立场之间的距离相等，所以他对两个候选人保持中立。因而，在$(x_D+x_R)/2$左边的摇摆选民更接近于x_D，所以会选择候选人D。同样，在$(x_D+x_R)/2$右边的选民会选择候选人R。①

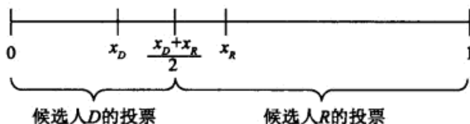


图6—7　竞选获胜者的产生

了解这些内容之后，我们就可以推断出立场是如何决定竞选结果的。若$x_D<x_R$，则候选人D得到的投票数量是$(x_D+x_R)/2$，因为$(x_D+x_R)/2$左边所有的选民会选择候选人R。因而，如果$(x_D+x_R)/2<\frac{1}{2}$，那么有不到一半的选民会给候选人D投票，候选人R获胜。相反，如果$\frac{1}{2}<(x_D+x_R)/2$，那么有超过一半的选民会给候选人D投票，候选人D获胜。若$x_R<x_D$，如果一个选民在两个候选人之间摇摆不定，那么之前的推理就适用。最后，$x_D=x_R$，则投票相同，因为候选人表明了类似的立场。

为了表述出候选人的收益函数，假定竞选成功的收益是2，失败的收益是0。如果候选人的选票相同，收益是多少呢？当两个候选人表明相同的立场或$(x_D+x_R)/2=\frac{1}{2}$时，才会出现上述的情况。那么在这种情形下，假定每个候选人获胜的概率相同，相关收益是1。

使用和描述立场是如何影响投票总数一样的分析方法，我们可以推出一个收益函数。若$x_D\leqslant x_R$，则候选人D的收益是：

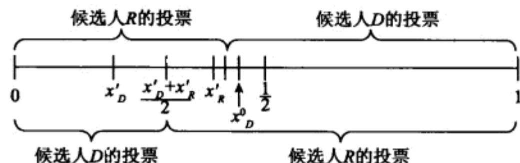
$$\begin{cases} 0, & \frac{x_D+x_R}{2}<\frac{1}{2} \text{ 且} x_D<x_R \\ 1, & \frac{x_D+x_R}{2}=\frac{1}{2} \text{ 或} x_D=x_R \\ 2, & \frac{1}{2}<\frac{x_D+x_R}{2} \text{ 且} x_D<x_R \end{cases}$$

若$x_R\leqslant x_D$，则：

$$\begin{cases} 2, & \frac{1}{2}<\frac{x_D+x_R}{2} \text{ 且} x_R<x_D \\ 0, & \frac{x_D+x_R}{2}=\frac{1}{2} \text{ 或} x_D=x_R \\ 1, & \frac{x_D+x_R}{2}<\frac{1}{2} \text{ 且} x_R<x_D \end{cases}$$

类似的，我们也可以求出候选人R的收益函数。

作为理解候选人动机的第一步，思考一下如图6—8所示的立场集合(x'_D,x'_R)。图6—8中的两个候选人都相对比较激进（两者都处于$\frac{1}{2}$的左边），但候选人D要比候选人R更激进，因为$x'_D<x'_R<\frac{1}{2}$。如图6—8底部所描述的，选民会同理想立场在$(x'_D+x'_R)/2$右边的所有选民一样投候选人R的票。这样候选人R就以超过半数的选票获胜。相反，如果候选人D的立场位于x'_R和$\frac{1}{2}$之间（即图中的x^0_D），那么她就会获得超过半数的选票，从而将她的收益从0增加至2，图6—8的顶端显示的就是这一结果。因而，(x'_D,x'_R)不是一个纳什均衡，因为候选人D可以通过放弃立场x'_D得到更高的收益。



一名候选人将立场定位于另一候选人立场和$\frac{1}{2}$之间会获得更高收益，候选人D选择立场x'_D时，选举结果如图底部所示，而选择立场x^0_D时的选举结果如图顶部所示。

图6—8　选择不同立场的选举结果

从更广泛的意义上来说，若候选人D选择不同于$\frac{1}{2}$的立场，则候选人R会通过选择$\frac{1}{2}$和他自身立场之间的立场以确保获胜。这就意味着候选人D肯定会失败。然而，一定会导致失败的策略显然不是候选人D的最佳策略，因为她至少可以选择和候选人R一样的立场和对手打平（得到的收益是1）。所以我们推断出立场组合中任何一方或双方的立场都不等于$\frac{1}{2}$时，这样的立场组合就不是一个纳什均衡。

除了策略集合$x_D=\frac{1}{2}=x_R$之外，我们已经剔除了所

①虽然位于$\frac{x_D+x_R}{2}$的投票人是中立的，但他投谁的票都无关紧要，因为有无限个投票人。——作者注

有的策略集合。在所剩的这一集合中，候选人平分选票，并且每人得到的收益是 1。假设候选人 D 选择了另一个立场 x'_D（如图 6—9 所示），这一选择导致她的选票数量从 $\frac{1}{2}$ 降至（$x'_D+\frac{1}{2}$）/2，从而使她的收益从 1 减少至 0。候选人 R 的情况与之类似，所以两个候选人的立场都为 $\frac{1}{2}$ 是唯一的纳什均衡。

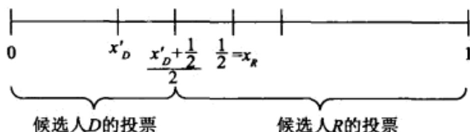


图 6—9　两名候选人都选择立场 $\frac{1}{2}$ 是唯一的纳什均衡

尽管只有两个候选人，但选举竞争会使选民无从选择，因为两个候选人的立场相同。候选人之所以会选择中间立场是因为该立场是中间选民最期待的立场：一半的选民比中间选民更激进，一半的选民比中间选民更保守。选择偏离中间选民立场的候选人就不会受选民欢迎，另一候选人就会利用这一点获得竞选胜利。这一结果被称为政策趋同，因为候选人会在相同的政策上趋于一致。

和一个简单模型通常出现的情况一样，这一分析的结果也是很极端的。实际上，候选人不会采纳相同或类似的立场。理论和现实的偏离至少是由于这一模型没有把一些重要的因素纳入其中。此外，我们假设所有的选民都投票，但事实上选民是否参加投票取决于自身的立场。一个候选人可能会采取极端立场以吸引具有相似立场的选民投票。尽管这一模型有很多不足，但它仍具有一定的借鉴意义，即要想获得选举，有一个推动力会使候选人在他们的立场上趋同。这一趋势即使不能完全实现政策趋同，但至少会有这样的趋向。

6.2 运用微积分求解纳什均衡（选读）

在本节中，我们会使用微积分来分析博弈中出现多重纳什均衡的情况，从而推导出纳什均衡的解。首先会对微积分的理论进行介绍，随后进入具体的案例分析。一些学生可能会比较喜欢这样的讲解顺序，可另外一些学生会觉得先从具体例子分析入手然后再接触较为抽象的理论对他们来说更为容易一些。我建议大家先通读这一部分的理论介绍，但是如果有人不明白的话也不用着急。这种情况下，可以先阅读第一个例子，然后返回到理论中重新阅读理解。

一个参与者的最优反应函数是指针对其他参与者所选策略采用一个使自身收益最大化的一个策略。在正式表述最优反应函数之前，先假设 $V_i(s_1,\cdots,s_n)$ 表示参与者 i 在 n 人博弈中的收益函数。在 $V_i(s_1,\cdots,s_n)$ 中插入一个策略组合后，就会出现一个数字，该数字代表参与者 i 赋予这个策略组合的值。参与者 i 对其他参与者所选策略（$s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n$）的最优反应是使 $V_i(s_1,\cdots,s_n)$ 最大化的一个策略。

一个参与者对其他参与者的某一特定策略组合只有一个最优反应，即只有一个策略能够产生最高收益。在本节分析的所有博弈中，参与者只有唯一的最优反应。令 $BR_i(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)$ 表示参与者 i 对其他 n-1 个参与者所选策略的唯一最优反应。假定 s_i 代表参与者 i 的策略组合，那么 BR_i 满足下列条件：①

$$V_i(s_1,\cdots,s_{i-1},BR_i,s_{i+1},\cdots,s_n)\geqslant V_i(s_1,\cdots,s_{i-1},s_i,s_{i+1},\cdots,s_n),\quad s_i\in S_i \qquad (6.1)$$

当每个参与者所选的策略是对其他参与者所选策略的一个最优反应时，策略组合（$s_1^*,\cdots,s_n^*$）是一个纳什均衡：

$$s_1^*=BR_1(s_2^*,\cdots,s_n^*)$$
$$s_2^*=BR_2(s_1^*,s_3^*,\cdots,s_n^*)$$
$$\vdots$$
$$s_n^*=BR_n(s_1^*,\cdots,s_{n-1}^*)$$

满足上述 n 个等式的策略组合就是纳什均衡。求纳什均衡的解就是就是求出这 n 个等式中的 n 个未知数，$s_1^*,\cdots,s_n^*$。可能没有解，可能有一个解，也可能有有限个解，甚至有可能有无限个解。

要运用这一求解方法，先必须用 BR_i 的求解不等式（6.1）推导出每个参与者的最优反应函数。假设 s_i=[0,1]，那么选择一个策略意味着从 [0,1] 区间选择一个数。进一步假设 V_i 的形式如图 6—10 所示。在图 6—10 中我们把参与者 i 之外的所有参与者的策略定位在：

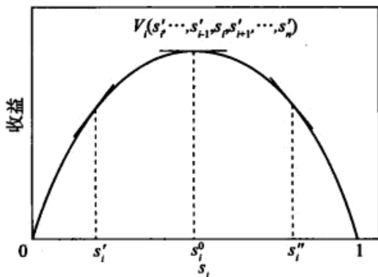
$$(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)=(s'_1,\cdots,s'_{i-1},s'_{i+1},\cdots,s'_n)$$

然后描述参与者 i 的收益是如何由 s_i 决定的。不难看出参与者 i 的最优反应是：

$$s_i^0=BR_i(s'_1,\cdots,s'_{i-1},s'_{i+1},\cdots,s'_n)$$

①回想 $s_i\in S_i$ 的含义是策略 s_i 是策略 S_i 中的要素。——作者注

图 6—10 中 V_i 的一个重要特性是它呈山形，并且 s^0_i 是它的最高点。收益函数在 $s_i=s^0_i$ 点的斜率是零，这一点可由平直切线看出。通常情况下，收益函数呈山形时，最佳策略就是收益函数斜率为零的那个点。在利用这一特性之前，需要先证明这一推断。



参与者 i 以外的所有参与者的策略定位是 $(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)=(s'_1,\cdots,s'_{i-1},s'_{i+1},\cdots,s'_n)$ 时，参与者 i 的收益与 s_i 的关系。

图 6—10　参与者 i 的收益与 s_i 的关系

设定在策略 s'_i 中，收益函数的斜率为正。斜率为正是指收益函数在 s_i 点呈上升趋势。因而，s'_i 不是一个最优反应，因为参与者 i 用一个数字较大的策略就可以得到更多收益。假定在策略 s''_i 中，收益函数的斜率为负（如图 6—10 所示）。既然策略 s''_i 在参与者 i 的策略中呈下降趋势，那么该策略不是最优反应，因为参与者 i 用一个数字较小的策略就可以得到更多收益。只有在斜率为零的一点上，收益才是最大的。

这样我们就可以引入莱布尼兹和牛顿在 17 世纪末提出的微积分学。微积分可以很容易地推导出函数斜率。该斜率就是函数的一阶微商，可以表示为：

$$\frac{\partial V_i(s_1,\cdots,s_n)}{\partial s_i}$$

按照之前的推理，参与者 i 的最优反应可以由使微商等于零的策略来表示，即：

$$\frac{\partial V_i(s'_1,\cdots,s'_{i-1},s^0_i,s'_{i+1},\cdots,s'_n)}{\partial s_i}=0。$$

总的来说，对其他参与者所选的任何策略，最优反应函数 $BR_i(s_1,\cdots,s_{i-1},s_{i+1},\cdots,s_n)$ 就是下面等式的解。

$$\frac{\partial V_i(s_1,\cdots,s_{i-1},BR_i,s_{i+1},\cdots,s_n)}{\partial s_i}=0$$

记住这一等式只有在参与者 i 的收益函数呈山形时才能确定最优反应。如果你不太懂微积分，那么可以帮助你判断收益函数呈山形的一个条件是收益函数的二阶微商总是负值，即：

$$\frac{\partial^2 V_i(s_1,\cdots,s_n)}{\partial s^2_i}<0$$

最后这一条件得到满足时，收益函数就被叫做严格凹函数。如果说这一概念对你来说有点太抽象，那么我们下面所举的两个例子应该会让你有一个更清晰的了解。

但在举例之前，需要提醒的是，在有无限个策略集合的情形下，纳什均衡肯定只有在某些条件下才会产生。设定一个参与者的策略集合是具有最低值和最高值的实数集合的一个区间，如 [0,10] 区间。另外，假设参与者的收益函数是圆滑的（曲线没有结），连续的（曲线没有出现猛增现象），以及呈山形（严格来说，这些曲线是凹函数）。满足前两个条件就可以确保我们探求的微商是存在的。根据以上假设，就会存在一个纳什均衡。而且如果博弈是对称的，那么就有一个对称纳什均衡。

注意这里描述的方法不适用于 6.1 小节中的博弈情形。要想求得微商，函数必须是可微函数，即该函数必须是圆滑和连续的。相同产品价格博弈情形中的收益函数不是连续的，因为一个商店和对手的价格持平时，会出现猛增现象。虽然价格一致承诺博弈中的收益函数没有出现猛增，但是当一个商店和另一个商店的定价相同时，会有一个结（参见图 6—6）。尽管如此，微积分仍然有助于求出这些博弈的解。我们会在本节中的最后一个例子（分析等额资助是如何影响慈善捐赠的）中证明这一论点。

案例研究

差别化产品的价格竞争

本章的第一个例子是商店销售相同产品。我们只是简单假设消费者仅仅关心价格，所以如果一个商店的定价较低，那么他们就会从那个商店购买。可是在多数市场情形中，商家会提供独特的产品。例如，你只能从可口可乐买到可口可乐，从麦当劳买到巨无霸（大号汉堡），从戴尔买到戴尔电脑。当然，你也可以从其他公司那里买到类似的产品，如七喜，庞然大物（一种汉堡）和惠普电脑等。然而，即使一个巨无霸比庞然大物要贵，一些人仍然会选择巨无霸，因为这一汉堡的口感独特且分量足。同样，也有人会更喜欢庞然大物，即使价格更贵。这些市场中的产品就具有差别化的特征。

现在就让我们来分析戴尔和惠普在电脑市场上的价格竞争。[①]鉴于上述的产品特性，假定戴尔电脑的需求(即它的销售量)由以下公式表示：

$$D^{Dell}(P^{Dell},P^{HP})=100-2P^{Dell}+P^{HP}$$

分析这一公式之后会发现，戴尔电脑的价格越高，销售量就越低。这一分析结果是有道理的，因为价格越高，一些消费者就会选择不买电脑或选择购买惠普电脑而不是戴尔电脑。而惠普的价格越高，戴尔的销售量就越大，因为潜在的惠普顾客会转而购买戴尔电脑。

戴尔的销售额等于它的销售量乘以它的定价，即：

$$P^{Dell}\times D^{Dell}(P^{Dell},P^{HP})=P^{Dell}\times(100-2P^{Dell}+P^{HP})$$

最后，假定一个戴尔电脑的生产和销售成本是 10 个单位。那么每台戴尔电脑的销售利润是 ($P^{Dell}-10$)，总利润（和收益）是：

$$V^{Dell}(P^{Dell},P^{HP})=(P^{Dell}-10)(100-2P^{Dell}+P^{HP}) \qquad (6.2)$$

假设惠普的成本更高，为 30 个单位，同理，我们可以推出它的总利润是：

$$V^{HP}(P^{Dell},P^{HP})=(P^{HP}-30)(100-2P^{HP}+P^{Dell}) \qquad (6.3)$$

该博弈的策略式如下：有两个参与者——戴尔和惠普，策略是价格。设公共的策略集合是 [0,100] 的区间数，收益函数如等式（6.2）和（6.3）所述。

为了求出纳什均衡的解，我们先推导出每个公司的最优反应函数。先分析戴尔的情况。图 6—11 表述的是如果 $P^{HP}=60$，戴尔的收益函数与自身定价之间的关系。从图 6—11 中可以发现戴尔的最优反应是 45，因为这一价格可以使戴尔的收益最大化（达到 2 450）。也可看出收益函数在 $P^{Dell}=45$ 这一点的斜率是零。和这一情形相反，戴尔的定价是 25 时，收益函数的斜率为正。也就是说，戴尔的收益呈上升趋势。譬如，P^{Dell} 的值从 25 升至 26 时，戴尔的收益会从 1 650 增至 1 728。因而，收益函数呈上升趋势不是一个最优反应。斜率为负的定价也不是一个最佳策略。例如，价格从 55 降至 54 时，戴尔的收益会从 2 250 增至 2 288。因此，最优反应是斜率为零的情况，这样调高或调低价格都不会产生一个更高收益。

从上述分析可以看出，戴尔的最优反应是收益函数的斜率为零所在点的价格。既然函数的斜率就是一阶微商，那么收益最大化的价格就是使一阶微商等于零的价格，即：

$$\frac{\partial V^{Dell}(P^{Dell},P^{HP})}{\partial P^{Dell}}=100-2P^{Dell}+P^{HP}-2P^{Dell}+20=0$$
$$=120-4P^{Dell}+P^{HP}=0$$

求出这个等式的解就可以推出戴尔的最优反应函数，即：

$$P^{Dell}=\frac{120+P^{HP}}{4} \text{ 或者 } BR^{Dell}=30+0.25P^{HP}$$

为了验证这一结论，把惠普电脑的价格替换为 60 后，你就可以推导出 $P^{Dell}=45$，如图 6—12 所示。

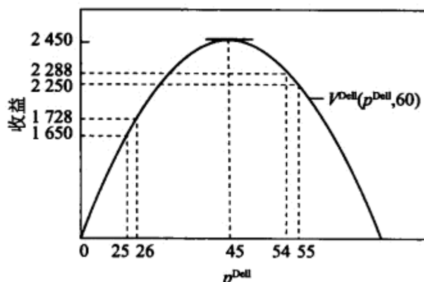


图 6—11　惠普定价为 60 时戴尔的收益函数

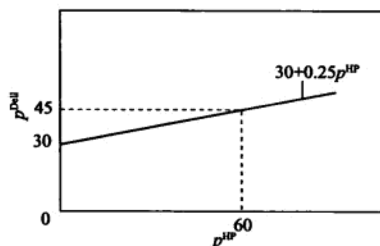


图 6—12　戴尔对惠普的最优反应函数

图 6—12 表述的是戴尔对惠普定价的最优反应函数。从图中可以看出和惠普的价格函数相比，戴尔的最优反应函数会逐步增加。换句话说，惠普的价格越高，购买戴尔的消费者越多，凭借拥有的更大市场需求戴尔可以制定更高的价格。因而，戴尔会从消费者身上赚取更多的利润，因为 P^{HP} 更高时消费者认为购买惠普电脑不划算。

按照相同的步骤，我们可以推出惠普的最优反应函数，即：

$$\frac{\partial V^{HP}(P^{Dell},P^{HP})}{\partial P^{HP}}=100-2P^{HP}+P^{Dell}-2P^{HP}+60=0$$

①为了准确模拟电脑市场，需要囊括所有的制造商，以及相关竞争对手，比如苹果电脑的Mac操作系统（不要和巨无霸Mac相混淆）。为了便于分析，我们只关注市场中的两家供应商。——作者注

$$=160-4P^{HP}+P^{Dell}=0$$

$$P^{HP}=\frac{160+P^{Dell}}{4}$$

$$BR^{HP}=40+0.25P^{Dell}$$

将代数的方法应用于这两个最优反应函数，就可以推出纳什均衡。价格组合（$\hat{P}^{Dell},\hat{P}^{HP}$）是一个纳什均衡，当两家电脑公司同时选择最优反应时：

$$\hat{P}^{Dell}=30+0.25\hat{P}^{HP}, \qquad (6.4)$$

$$\hat{P}^{HP}=40+0.25\hat{P}^{Dell}. \qquad (6.5)$$

现在有两个等式和两个未知数，只要找出满足这两个等式的价格组合即可。我们求出的这个价格组合就是图 6—13 所示的两个最优反应函数的交叉点。在这样一个价格组合交叉点上，每家公司的价格都是最优反应。

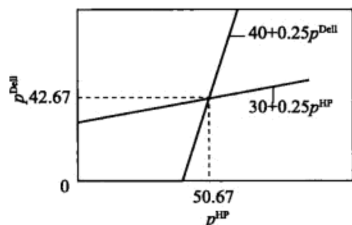


图 6—13　纳什均衡是两个最优反应函数的交叉点

按照下面的代数方法，我们可以求出联立方程（6.4）和(6.5)的解。用方程式(6.5)右边的项替换方程式(6.4)中的$\hat{P}^{HP}$ 可以得出：

$$\hat{P}^{Dell}=30+0.25\ (40+0.25\hat{P}^{Dell})$$

这样，就只剩下一个等式和一个未知数。接下来，再用代数的方法可以得出：

$$\hat{P}^{Dell}=30+10+0.062\,5\,\hat{P}^{Dell}$$

$$\hat{P}^{Dell}(1-0.062\,5)=40$$

$$\hat{P}^{Dell}=42.67$$

这样，一台戴尔电脑的均衡价格就是 42.67。将这一价格代入惠普的最优反应函数，就可以得到惠普的均衡价格，即：

$$\hat{P}^{HP}=40+0.25\times 42.67$$

$$\hat{P}^{HP}=50.67$$

因为联立方程（6.4）和（6.5）只有一个解，所以我们可以求出唯一的纳什均衡，即戴尔电脑的定价为 42.67，惠普电脑的定价为 50.67。看到这样的结论，我想连牛顿和莱布尼兹都会向我们致敬吧。

6.2 理解测验

假设戴尔和惠普电脑的最低成本是 20。则戴尔的收益函数是：

$V^{Dell}(P^{Dell},P^{HP})=(P^{Dell}-20)(100-2P^{Dell}+P^{HP})$，

而惠普的收益函数是：

$V^{HP}(P^{Dell},P^{HP})=(P^{HP}-20)(100-2P^{HP}+P^{Dell})$

求出所有的纳什均衡。

公地悲剧——猛犸象的灭绝

更新世时期末，美洲大陆一半以上的大型哺乳动物大量灭绝。在试图解释这一现象的众多假设中，认为是由人为捕猎造成的观点颇为引人注目。有证据显示美洲大陆出现大量人口的时期可以追溯到 13 400 年前，而那之后仅仅过了 1 200 年，动物大规模灭绝的历史又重演了。



捕杀猛犸象

最近进行的一项关于原始人类和环境之间相互关系的计算机模拟实验证明上述假设是成立的。其中的一个模拟结果如图 6—14 所示，图中的粗黑线代表人口数量，而其他每条线代表被人类捕杀的一个物种。从图 6—14 中可以看出，大多数物种的规模都趋于零（代表物种灭绝），而且从人类初现到这些物种的灭绝之间的时间只有 1 299 年，这和有关证据显示的结果十分接近。

要探究人类是如何捕杀其他物种并导致其灭绝的，我们不妨先回到冰河时期，看看原始人类的捕杀产生了怎样的后果。假设有 n 个捕猎者，每个猎手决定付出多大的努力才能成功捕获猛犸象。e_i 代表猎手 i 的努力，并假定 $e_i \geq 0$，这样策略组合包括所有非负（实）数。（我们可以将努力理解为花费的时间量，在这种情况下给这个策略组合设定上界是合适的。）

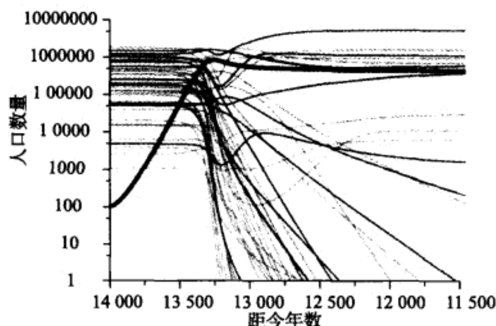


图 6—14　人口数量规模（粗黑线）与珍稀物种（浅灰线）

被捕杀的猛犸象总数取决于所有猎手付出的努力。设 $E=e_1+e_2+\cdots+e_n$ 为所有猎手的整体努力，则可以推出被捕杀的猛犸象总数（以重量为单位）是 $E(1\,000-E)$，如图 6—15 所示。从图中可以看出 E 随着整体努力的变化先增加后减少，这是因为有两股力量在起作用。对于既定数量的猛犸象，付出的努力越多意味着死亡的猛犸象也越多。然而，因为猛犸象越少繁衍的后代也越少，所以付出的努力越多会导致可捕杀的猛犸象数量越少。当整体努力程度非常低（$E<500$）时，稍稍努力一点就可以捕获更多的猛犸象，这样第一股力量比第二股力量强大。当整体努力的程度非常高（$E>500$）时，付出的努力越多可以捕杀的猛犸象就越少，所以第二股力量占上风。超过 500 的整体努力就属于过度捕杀，即人类捕杀猛犸象的速度超过了猛犸象繁衍生息所需的时间。

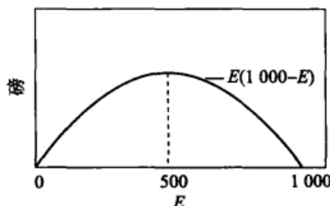


图 6—15　捕杀猛犸象付出的努力和捕获的猛犸象之间的关系

我们已经分析了人类的整体努力是如何决定被捕杀的猛犸象总数的。可还有一个问题是捕获的猛犸象是如何在猎手之间分配的。平均主义的方法是允许每个猎手获得相同份额的猎物，即 $E(1\,000-E)/n$。也许会有人说更公平的办法是谁付出的努力更多就有权得到更多的猎物。这一办法可以由部落首领的评判实现，但前提是他观察了所有猎手付出的努力。也可以假设猎手们单独捕猎（或以小组为单位进行捕猎），然后分得他们实际捕获的猛犸象，这样一来谁付出的努力越多，分得的猎物就越多。

使用这个平均分配的方法，我们求出猎手 i 得到的猎物份额应该和他的努力在整体努力中所占的份额相等，即：

$$\left(\frac{e_i}{E}\right)E(1\,000-E)=e_i(1\,000-E)$$

整体努力是 $E(1\,000-E)$，猎手 i 在当中的份额是 e_i/E。最后，需要剔除猎手 i 努力的成本，假定为 $100e_i$。那么一个猎手的收益函数就是：

$$V_i(e_1,\cdots,e_n)=e_i[1\,000-(e_1+e_2+\cdots+e_n)]-100e_i$$

设其他 n−1 个猎手的努力为既定项，我们就可以描述出如图 6—16 所示的猎手 i 的收益函数。注意该收益函数呈山形。①

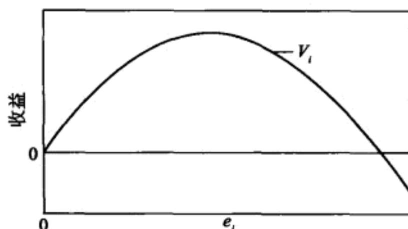


图 6—16　猎手 i 的收益函数

构建起该博弈模型后，接下来就是求出一个猎手的最优反应函数。要做到这一点，首先需要求出和他努力相关的收益函数的微商并将其设定为零，即：

$$\frac{\partial V_i(e_1,\cdots,e_n)}{\partial e_i}=1\,000-(e_1+e_2+\cdots+e_n)-e_i-100=0 \quad (6.6)$$
$$=900-(e_1+e_2+\cdots+e_{i-1}+e_{i+1}+\cdots+e_n)-2e_i$$
$$=0$$

接着求出等式（6.6）中的 e_i，从而得出最优反应函数：

$$BR_i=450-\left(\frac{1}{2}\right)(e_1+e_2+\cdots+e_{i-1}+e_{i+1}+\cdots+e_n)$$

一个纳什均衡是 n 种努力程度，这些努力使 n 个等式成立，从而实现每个猎手收益的最大化，即：

$$e^*_1=450-\left(\frac{1}{2}\right)(e^*_2+\cdots+e^*_n)$$

$$e^*_2=450-\left(\frac{1}{2}\right)(e^*_1+e^*_3+\cdots+e^*_n)$$

①通过仔细观察也可以发现 $\partial^2 V_i(e_1,\cdots,e_n)/\partial e^2_i=-2<0$。——作者注

$$\vdots$$

$$e^*_n=450-\left(\frac{1}{2}\right)(e^*_1+\cdots+e^*_{n-1}) \qquad (6.7)$$

注意所有猎手的最优反应函数形式相同，即 450 减去其他所有猎手整体努力的一半。因为该博弈是对称的，所以结果并不令人意外。下面求出一个对称纳什均衡以简化我们的工作。

一个对称纳什均衡是平均努力程度，设为 e^*，如果其他 $n-1$ 个猎手选择 e^*，那么猎手 i 也选择 e^* 是最佳策略。用 e^* 替换等式（6.7）中的 $e^*_1,\cdots,e^*_n$，就可以得出 n 个相同等式，即：

$$e^*=450-\left(\frac{1}{2}\right)(n-1)\,e^*$$

现在只需求出只有一个未知数的等式，即：

$$e^*=450-\left(\frac{1}{2}\right)(n-1)\,e^*$$

$$e^*\left[1+\left(\frac{1}{2}\right)(n-1)\right]=450$$

$$e^*\left(\frac{2+n-1}{2}\right)=450$$

$$e^*=\frac{900}{n+1}$$

在一个纳什均衡中，每个猎手付出的努力都等于 $900/(n+1)$。如果有 9 个猎手，那么每个猎手选择的努力程度是 90。猎手数量越多每个猎手努力的程度越小。所以，如果有 10 个而不是 9 个猎手的话，每个猎手的努力程度就只有 82。因为有更多的猎手追赶同一群猛犸象，所以每个猎手的努力程度就更小。这样一个猎手就发现捕杀小型动物如兔子或收菜得到的实惠更多。

现在分析一下所有猎手的整体努力程度，用 ne^* 来表示，且 $ne^*=900n/(n+1)$。图 6—17 表述的就是这一努力程度。从图中可以看出，整体努力程度会随着猎手数量的增加而增强。[①]虽然猎手更多时每个猎手捕获的更少，但其他猎手的努力程度会抵消这一影响，所以捕猎的整体努力程度是呈上升趋势的，而且出现了过度捕杀猛犸象种群的状况，也就是说，猎手捕获的猛犸象总数超过了它的种群所能提供的数量。当猎手的整体努力程度为 500 时，猛犸象所能提供的种群数量达到最大（参见图 6—15）。可是，至少有两个猎手时，整体的均衡努力就会超出这一数量（猛犸象总数），即：

$$\frac{900n}{n+1}>500$$

$$900n>500n+500$$

$$400n>500$$

$$n>1.2$$

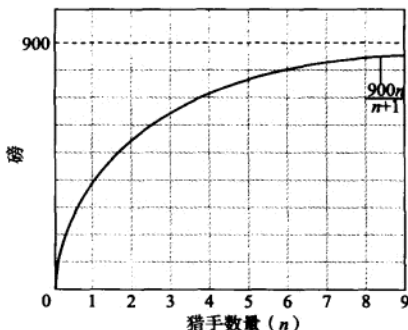


图 6—17　所有猎手选择均衡努力程度时猎手总数与集体努力程度之间的关系

猛犸象遭到人类的过度捕杀，这一现象被生态学家加勒特·哈丁（Garrett Hardin）称为公地悲剧中的一个范例。**公地悲剧是指两个或更多的人过度使用一项资源，从而导致整体的利益受损。**过度捕捞智利海洋中的鱼类，过度砍伐亚马孙丛林中的树木，过快开采公共水库中的石油都是公地悲剧的典型例子。参与者之间的相互依存性（经济学家称之为外在性）是问题的核心所在。一个猎手捕杀猛犸象时，并没有考虑他的行为对其他猎手的整体利益产生的负面效应（即他们可以捕杀的猛犸象更少了）。因此，从整个人类群体的角度来说，每个猎手捕杀了太多的猛犸象。

毫无疑问，当前公地悲剧的一个最突出的例子就是全球气候变化。美国环境保护局称：“从工业革命开始，大气中的二氧化碳含量增加了近 30%；甲烷含量增加了两倍多；一氧化氮的含量也增加了 15%。”同一时期，地球表面的平均温度增加了华氏 0.5 度至 1 度，海平面上升了 10 至 20 厘米。人们对这些客观事实没有异议，但在是否是气候变化导致了地球温度的升高这一方面还存在分歧。如果原因的确是这样，那么让这一公地悲剧不再重演的唯一方法就是《京都议定书》（*Kyoto Accord*）所提出的通过人类共同的努力以减少温室气体的排放。

①这一情形可以由下面的公式表示，即$\frac{\partial ne^*}{\partial n}=\frac{900(n+1)-900n}{(n+1)^2}=\frac{900}{(n+1)^2}>0$。——作者注

配额资助对慈善捐助的影响力

在最后这个例子中，收益函数不是呈山形的，甚至也不是连续的。换句话说，在之前的两个例子中使用的方法在这里不适用。而我之所以举这个例子是因为以下两点原因：

（1）这个例子会提醒你不要不加分析地运用在6.2节伊始提出的基于微积分的方法。因为使用该方法必须满足收益函数是可微函数（即函数是连续的并且当中没有结）而且呈山形，所以你必须确定函数具备了这样的特性后才能加以运用。

（2）即使不能使用上述方法，微积分仍可以帮助我们求出参与者的最优反应函数。

假定一个慈善家打算为他中意的慈善项目捐款300万美元。虽然他很富有，但他一个人还是无法捐助如此大的一笔款项。所以，为了鼓励其他人也捐款，他提出了一个配额资助方案，即如果从其他捐款人那里筹集到200万美元，他就捐助100万美元。但如果筹集到的数目不到200万美元，那么他就不捐款。其实这一方案并不新颖，因为许多的慈善机构和非盈利组织就采用这一方案。事实上，美国国家公共电台在筹集资金的活动中就经常使用与之类似的方案。博弈理论将揭示出配额资助是如何吸引更多捐款的。

假设有10个潜在的捐款人同时决定捐助多少款项。s_i代表捐款人i的捐款，s_{-i}是指捐款人i之外的所有捐款，即：

$$s_{-i}=s_1+\cdots+s_{i-1}+s_{i+1}+\cdots+s_{10}$$

假设一个捐款人的策略集合是0到500 000的区间数，单位为美元。捐款人i的收益表述为：

$$\left(\frac{1}{5}\right)(s_{-i}+s_i)-s_i$$

其中，$\left(\frac{1}{5}\right)(s_{-i}+s_i)$是指从用于崇高事业的钱款中得到的利益且该利益取决于总的捐款数；$-s_i$是指捐助导致的个人损失。

如果没有配额资助，会有人捐款吗？在没有配额资助的条件下，捐款人1的收益函数是：

$$V_1(s_1,\cdots,s_n)=\left(\frac{1}{5}\right)(s_1+\cdots+s_{10})-s_1$$

这表示捐款人策略的收益函数并不呈山形。实际上，情形要比这简单得多。求出关于s_1的$V_1(s_1,\cdots,s_n)$中的一阶微商：

$$\frac{\partial V_1(s_1,\cdots,s_n)}{\partial s_1}=-\frac{4}{5}$$

我们就能推出捐款人的收益随着她的捐款而减少。她每捐一美元，遭受的损失就是1美元，而得到的利益只有20美分。所以她每捐一美元，收益就会减少80美分，即$\frac{4}{5}$美元。这样，她捐得越多，就感觉越郁闷。因而，不捐款是最佳策略。不论其他人是否捐款，不参加捐款都是最佳策略，因此，不捐款是一个占优策略。最后，因为捐款人1和其他捐款人一样，所以他们选择不捐款也是一个占优策略。所有10个捐款人都不捐款的策略组合就是唯一的纳什均衡。看来我们的捐款活动一开始就出现了坏兆头。

现在假设有等额资助，那么捐款人1的收益函数就是：

$$V_1(s_1,\cdots,s_n)=\begin{cases}\left(\frac{1}{5}\right)(s_1+\cdots+s_{10})-s_1 & s_1+\cdots+s_{10}<2\,000\,000\\ \left(\frac{1}{5}\right)(s_1+\cdots+s_{10}+1\,000\,000)-s_1 & 2\,000\,000\leqslant s_1+\cdots+s_{10}\end{cases}$$

若总捐款不足2 000 000，则捐款人1的收益和没有配额资助条件下的收益相同。然而，若总捐款达到2 000 000，则每个捐款人的收益猛增$\left(\frac{1}{5}\right)\times 1\,000\,000$或200 000。出现这一数量猛增，使得收益函数不具有连续性，进而也就不是可微函数，所以我们无法从求出微商着手。不过，更细心一点的话，就会发现微积分仍可以帮助我们解决问题。

先求出一个捐款人的最优反应函数（因为博弈是对称的，所以捐款人的最优反应相同）。以捐款人1为例，设s_{-1}（$=s_2+\cdots+s_{10}$）为其他9个捐款人的总捐款数。仔细观察如果$s_{-1}\geqslant 2\,000\,000$，那么无论捐款人1是否捐款都会出现配额资助，所以，无论s_1取何值，捐款人1的收益是：

$$\left(\frac{1}{5}\right)(s_1+s_{-1}+1\,000\,000)-s_1$$

这个关于s_1的表达式的微商是$-\frac{4}{5}$，所以捐款人1的收益会因为她捐款而减少。因而，最优的捐款数是最低可行的捐款数，即零。

假定$s_{-1}<2\,000\,000$，这时其他捐款人的捐款总数就不会产生配额资助。只要捐款人1的捐款数不能使总捐款数达到2 000 000，即$s_1+s_{-1}<2\,000\,000$，那么捐款人1的收益就是：

$$\left(\frac{1}{5}\right)(s_1+s_{-1})-s_1$$

这一表达式的微商是$-\frac{4}{5}$，因而她的收益肯定会因为捐款而减少。这样一来，如果捐款人 1 的捐款数不能产生等额资助的话，那么她就不应该捐款。其次，假设她的捐款数能产生配额资助，即 $s_1+s_{-1} \geqslant 2\,000\,000$，那么她这时的收益为：

$$\left(\frac{1}{5}\right)(s_1+s_{-1}+1\,000\,000)-s_1$$

同样，求出的微商也是$-\frac{4}{5}$，所以她的捐款数越少得到的收益就越多。因而，如果满足出现配额资助的条件，那么捐款人 1 会认为最佳策略是尽可能使捐款数降至最低，即 $2\,000\,000-s_{-1}$。

总之，若 $s_{-1}<2\,000\,000$，则捐款人 1 的最优反应是不捐款或是使总捐款达到 2 000 000 的最小捐款。举几个例子应该可以更好地解释这一论点背后的思维逻辑。假设 $s_{-1}=1\,900\,000$，由此得出的捐款人 1 的收益函数如图 6—18 所示。若 $s_1<100\,000$，则总的捐款数少于 2 000 000 的目标，在这样的总捐款数中，捐款人 1 的收益是：

$$\left(\frac{1}{5}\right)(s_1+1\,900\,000)-s_1=380\,000-\left(\frac{4}{5}\right)s_1$$

当 s_1 达到 100 000 时，她的收益会因为出现 1 000 000 配额资助而猛增至 500 000，即：

$$\left(\frac{1}{5}\right)(100\,000+1\,900\,000+1\,000\,000)-100\,000$$
$$=600\,000-100\,000=500\,000$$

当 $s_1>100\,000$ 时，捐款人 1 的收益又一次减少，即：

$$\left(\frac{1}{5}\right)(s_1+1\,900\,000+1\,000\,000)-s_1=580\,000-\left(\frac{4}{5}\right)s_1$$

正如图 6—18 显示的那样，捐款人 1 的最佳捐款是 100 000，也就是使配额资助出现的最小捐款。

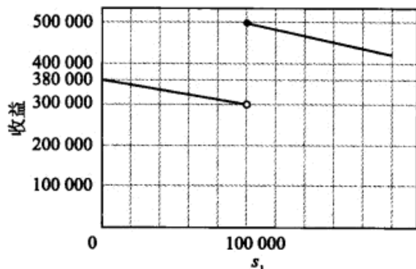


图 6—18　s_{-1}=1 900 000 时，捐款人 1 的收益函数

相反，假定 $s_{-1}=1\,650\,000$。在这种情形下，捐款人需要捐出 350 000 才可以产生等额资助。如图 6—19 所示，捐款人 1 这时会选择不捐款。和之前一样，她的收益只有在总捐款数达到 2 000 000 时才不会减少。当总捐款数为 2 000 000 时，她的收益从 50 000 增至 250 000，之后又出现减少的情况。捐款人 1 的捐款数为零时可以使她的收益最大化。至少捐款 350 000 才可以产生等额资助的捐款数目对于任何一个捐款人来说都有些太大了。

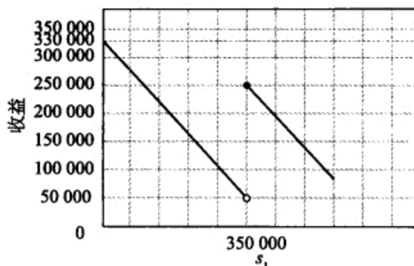


图 6—19　s_{-1}=1 650 000 时捐款人 1 的收益，她的最优捐款数额为零

其他人的总捐款数达不到出现配额资助的条件时，一个捐款人的最佳捐款就是零或可以产生配额资助的捐款。下面就对比这两种选择以确定哪一个更优。假使 $s_{-1}<2\,000\,000$，当：

$$\left(\frac{1}{5}\right)3\,000\,000-(2\,000\,000-s_{-1})\geqslant\left(\frac{1}{5}\right)s_{-1}$$

时，我们推断出捐款人 1 会选择捐款而不是不捐款，这样总捐款数就可以达到 2 000 000。上述不等式的左边是从捐款 $2\,000\,000-s_{-1}$ 中得到的收益，而右边是不捐款得到的收益。求出这一不等式中 s_{-1} 的解，我们得出：

$$\left(\frac{4}{5}\right)s_{-1}\geqslant 1\,400\,000\text{ 或 }s_{-1}\geqslant 1\,750\,000$$

因此，若 $s_{-1}>1\,750\,000$，则捐款人 1 的最佳捐款为 $2\,000\,000-s_{-1}$，从而确保会出现配额资助。若 $s_{-1}<1\,750\,000$，则捐款人 1 会选择不捐款。当 $s_{-1}=1\,750\,000$ 时，这两个选择对她来说是一样的（我们推测她会捐款 250 000）。

由于对称性，这一结论适用于所有的捐款人。因而，捐款人 i 的最优反应函数是：

$$BRi=\begin{cases}0 & s_{-i}<1\,750\,000\\ 2\,000\,000-s_{-i} & 1\,750\,000\leqslant s_{-i}<2\,000\,000\\ 0 & 2\,000\,000\geqslant s_{-i}\end{cases}$$

图 6—20 描述的就是这一函数。

现在我们用最优反应函数求出对称纳什均衡。我们需要推导出其他 9 个捐款人选择的捐款数对一个捐款人来说是一个最优捐款的捐款数。图 6—20 表述的是捐款数为零的一个对称纳什均衡。若其他 9 个捐款人每人的

捐款数为零，则 $s_{-1}=0$，根据捐款人 i 的最优反应函数，他的最优捐款数也是零。

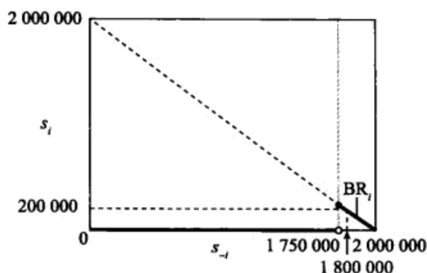


图 6—20　捐款人 i 的最优反应函数

捐款人的捐款数不为零的策略组合是一个纳什均衡吗？如果一个捐款人捐款，她的最优捐款是可以产生配额资助 2 000 000 的最低捐款数。在有 10 个捐款人和一个对称策略组合的条件下，可以求出每个人会捐款 20 000。这样，s_{-1}=1 800 000（因为其他捐款人每人捐款 200 000），从图 6-20 可以看出，一个捐款人也会捐款 200 000 作为他的最佳策略。因而，每个人捐款 200 000 的策略组合是一个纳什均衡，从而可以确保出现配额资助。

在没有配额资助的条件下，唯一的均衡是每人都不捐款。所以，如果至少筹集 2 000 000 捐款时，捐款 1 000 000 就是一个纳什均衡，因为此时捐款数量会达到 2 000 000，从而出现配额资助。配额资助的作用是增大一笔捐助的边际影响力。假定其他捐款人的捐款总额是 180 000，那么捐款 200 000 的捐款人实际上使捐款总数增加了 1 200 000，所以配额资助会诱使她捐款。因为所有的捐款人都会做出这样的推理，所以他们每个人都会捐款，从而产生配额资助。

本章总结与练习题

本章总结

本章研究了出现**连续策略集合**（由实数区间表示）的博弈情形。在有无限个策略组合的情形下，用穷举搜索的方法求出纳什均衡是不太可行的。所以，在 6.1 小节中，我们阐释了如何通过推断参与者的动机来剔除许多潜在的纳什均衡策略组合。在相同产品价格竞争的博弈中，当产品价格超过了成本时，每个公司都有稍稍压低对手价格的动机。因为只要商店定价高于其成本，压价的动机就会出现，所以所有高于成本的定价策略组合都不是纳什均衡。利用这一理念，我们可以剔除许多可能成为均衡的策略组合，从而最终推出商店以成本定价的策略组合是纳什均衡的结论。

在 6.2 节当中，我们介绍了一个利用微积分来求出纳什均衡的方法。一个参与者的收益函数是可微函数（连续的并且没有结）而且呈山形时，他的最优反应是令其（策略）收益函数的一阶微商为零的策略。如果微商为正值（负值），那么参与者可以通过增加（减少）他的策略值来增加他的收益。只有当微商为零时的策略组合才能使参与者的收益最大化。这一认识使我们得到一个等式，该等式可以很容易地求出一个参与者的最优反应函数。所有参与者的最优反应函数可以用来求出一个纳什均衡。

运用上述基于微积分的方法，我们求出了两种博弈情形中的唯一纳什均衡。第一种是公司提供差别化产品，并进行价格竞争的博弈情形。第二种是原始人类如何分配劳力用于捕猎的博弈情形。分析慈善捐款的例子是为了提醒我们如果要运用基于微积分的方法，首先必须确定收益函数是可微函数并且呈山形。

纳什均衡的一个共性是他们在所有的策略组合集合中都不是收益占优的。也就是说，和一个纳什均衡相比，如果所有的参与者以某种方式改变了他们的策略选择，那么他们都可以得到更高的收益。这是因为每个参与者在最大化自身的收益时，都会忽略她所选策略对其他参与者收益所产生的影响。参与者不会从集体利益最优的角度采取行动，正如公地悲剧所展现的那样。从整体的角度来说，公共资源被过度开发（如过度捕杀动物致使其灭绝，公路拥堵导致交通事故，温室气体污染环境等），解决这一问题的对策是参与者相互配合、相互制约以减少对资源的使用量。尽管相互协调的对策与本章所分析的博弈中的纳什均衡并不一致，但是我们会在 13 至 15 章分析更系统和严谨的理论公式是如何实现这一协调的。

PDG

练习题

练习 7 到 12 需要使用微积分的方法。

1. 一个博弈理论家在沿街散步时发现了一张 20 美元的钞票。正当他捡起钞票时，住在同一街区的两个小孩，简和蒂姆跑了过来，问他是否可以将这张钞票给他们。因为博弈理论家天生就慷慨大方，所以他很愿意让他们两个得到这张钞票，但有一个前提条件，即每人递交一份想从 20 美元中分得份额的申请，t 代表蒂姆想分得的份额，j 代表简想分得的份额。j 和 t 为 [0,20] 的区间数。若 $j+t \leq 20$，则两个人可以得到想要的份额，剩下的 $20-j-t$ 在他们之间平分。可是若 $j+t>20$，那么他们将一无所获，这 20 美元归这个博弈理论家所有。简和蒂姆是这一博弈的参与者，假设他们每人的收益等于他们得到的钱数。求出所有的纳什均衡。

2. 在一些总统选举中，会出现一个第三方党派候选人。在 1992 年，罗斯·佩罗（Ross Perot）作为一个独立候选人与民主党候选人比尔·克林顿和共和党候选人乔治·布什展开竞争，并获得了 19% 的选票。近来作为 2000 年和 2004 年绿党候选人的拉尔夫·纳德（Ralpa Nader）也进行了第三方总统竞选的尝试。鉴于这些情况，分析一个有三方候选人（分别用 D、R、I 表示）的竞选模式。和 6.1 小节中描述的情形相同，每个候选人从 [0,1] 区间选择一个立场。若一个候选人得到的选票比其他两个候选人的多（这样他就肯定获选）则他得到的收益是 1；若与另一个候选人得到同样的多数选票，则得到的收益是 $\frac{1}{2}$；若三人平分选票，则收益是 $\frac{1}{3}$。若得到的选票少于另一个候选人（这意味他肯定落选），则得到的收益是 0。假定投票人会选择和自己的立场最接近的候选人。三个候选人的立场都处于 $\frac{1}{2}$ 点时的策略集合是一个纳什均衡吗？

3. 回到 6.2 小节中的相同产品价格竞争博弈。假设商店 2 的成本是 15，商店 1 的成本仍是 10。假定如果两家商店定价相同，那么所有的消费者都会从商店 1 购买。当商店 2 的收益函数是：

$$\begin{cases} 0 & p_1 \leq p_2 \\ (p_2-15)(100-p_2) & p_2 < p_1 \end{cases}$$

时，商店 1 的收益函数是：

$$\begin{cases} (p_1-10)(100-p_1) & p_1 \leq p_2 \\ 0 & p_2 < p_1 \end{cases}$$

写出所有的纳什均衡。

4. 回到 6.1 小节中的价格一致承诺博弈中。

a. 假设两家商店定价相同并且价格都超过了 55，这种情形是一个纳什均衡吗？

b. 假设两家商店定价相同并且价格都不到 10，这种情形是一个纳什均衡吗？

c. 写出所有的非劣对称纳什均衡。

5. 两个制造商（用 1，2 表示）争夺 100 个相同的顾客。每个制造商以价格和产品质量作为他的策略选择，每个变量可以是任意非负的实数。p_i 和 x_i 分别表示制造商 i 的产品价格和质量。制造商每件产品的生产成本是 $10+5x_i$。注意在这一表达式中质量越好成本越高。若制造商 i 将产品卖给 q_i 个消费者，则它的总成本是 $q_i(10+5x_i)$。每个顾客选择提供最高价值的制造商，其中购买制造商 i 产品的价值是 $1\,000+x_i-p_i$，更好的质量和更低的价格意味着更高的价值。一个制造商的收益是它的利润，该利润等于 $q_i(p_i-10-5x_i)$。若一个制造商提供的价值更高，则所有的 100 个顾客都选择从它那里购买。若两个制造商提供的价值相同，则他们各得到 50 个顾客。求出所有的纳什均衡。

6. 求出 6.2 小节中所讲的慈善捐款博弈中所有的纳什均衡。

7. 在一个两人博弈中，参与者 1 的收益函数是：

$$V_i(x_1,x_2)=x_1+10x_1x_2$$

而参与者 2 的收益函数是：

$$V_2(x_1,x_2)=x_2+20x_1x_2$$

参与者 1 的策略集合是 [0,100] 的区间数，参与者 2 的策略集合是 [0,50] 的区间数。写出所有的纳什均衡。

8. 武器集结被认为是导致第一次世界大战的诱因之一。德国和英国之间的海军军备竞赛尤为明显。在 1889 年，英国采取了维持海上霸主地位的政策，要求自己的海军规模必须是第二大海军规模的两倍半。这种激进的姿态促使德国扩大自己的海军规模，而这种做法又使得英国进一步壮大自己的海军规模，两国的竞争如此进行下去。虽然在 1899 和 1907 年进行过裁军，但军事竞赛仍在继续。1914 年爆发第一次世界大战时，英国的海军吨位是 2 205 000 磅（略小于对手两倍半的规模），而作为第二大海军的德国海军吨位是 1 019 000 磅。根据上述

情况，我们设想一个两国（分别用 1 和 2 表示）军备竞赛博弈情形。国家 i 的军费开支用 x_i 表示，并限制在 [1,25] 内。增加军费开支的利益是增强国家安全或发动战争的能力，二者取决于相对的军费开支。这样，假定国家 1 的利益是 $36\left(\frac{x_1}{x_1+x_2}\right)$，则该利益会随着国家 1 的开支在总开支中的比例而增加，损失仅仅是 x_1，因而国家 1 的收益函数是：

$$V_1(x_1,x_2)=36\left(\frac{x_1}{x_1+x_2}\right)-x_1$$

同理，国家 2 的收益函数是：

$$V_2(x_1,x_2)=36\left(\frac{x_2}{x_1+x_2}\right)-x_2$$

这些收益函数呈山形。

a. 求出每个国家的最优反应函数。

b. 写出一个对称纳什均衡。

9. 参与者 1 和参与者 2 进行一场博弈，其中参与者 i 的策略用 z_i 表示，并且是任意一个非负实数。参与者 1 的收益函数是：

$$V_1(z_1,z_2)=(100-z_1-z_2)z_1$$

而参与者 2 的收益函数是：

$$V_2(z_1,z_2)=(80-z_1-z_2)z_2$$

这些收益函数呈山形。求出所有的纳什均衡。

10. 一对夫妻的结婚纪念日即将来临，他们每个人正在决定庆祝的花销。g_H 代表丈夫给妻子的花销，g_W 代表妻子给丈夫的花销。假定他们同意每个人给对方的花销最多是 500。一个参与者的策略集合就是属于 [0,500] 的区间数。每一方配偶都想送个更大的礼物，但是又不想花钱。由此，丈夫的收益函数可以表示为：

$$V_H(g_H,g_W)=50g_H+\left(\frac{1}{4}\right)g_Hg_W-\left(\frac{1}{2}\right)(g_H)^2$$

在这一收益函数中，交换礼物得到的利益是 $50g_H+\left(\frac{1}{4}\right)g_Hg_W$。因为男人送的礼物通常更大，所以这一利益随着妻子礼物的大小而增加：

$$\frac{\partial[50g_H+\left(\frac{1}{4}\right)g_Hg_W]}{\partial g_W}=\left(\frac{1}{4}\right)g_H>0$$

丈夫从收到妻子礼物中得到的利益由 $50g_H+\left(\frac{1}{4}\right)g_Hg_W$ 项表示，且这一利益随着他的礼物的大小而增加，即：

$$\frac{\partial[50g_H+\left(\frac{1}{4}\right)g_Hg_W]}{\partial g_H}=50+\left(\frac{1}{4}\right)g_W>0$$

可是有利益就有损失。丈夫买一个 g_H 大小的礼物受到的损失可以由他收益函数中的 $-g_H\times g_H$ 项表示。因而，从利益中减去他受到的损失，就可以得出之前描述的丈夫的收益函数。同理，妻子的收益函数为：

$$V_W(g_H,g_W)=50g_w+2g_Hg_W-\left(\frac{1}{2}\right)(g_W)^2$$

这些收益函数都呈山形。

a. 求出每一方配偶的最优反应函数并绘图表示。

b. 写出一个纳什均衡。

c. 假定丈夫的收益函数和妻子的收益函数相同，即都是：

$$V_H(g_H,g_W)=50g_H+2g_Hg_W-\left(\frac{1}{2}\right)(g_H)^2$$

求出一个纳什均衡。（提示：别忘了策略集合。）

11. 参与者 1,2 和 3 正进行一场博弈，当中参与者 i 的策略用 x_i 表示，x_i 为任一非负实数。参与者 1 的收益函数是：

$$V_1(x_1,x_2,x_3)=x_1x_2x_3-\frac{1}{2}(x_1)^2$$

参与者 2 的收益函数是：

$$V_2(x_1,x_2,x_3)=x_1x_2x_3-\frac{1}{2}(x_2)^2$$

参与者 3 的收益函数是：

$$V_3(x_1,x_2,x_3)=x_1x_2x_3-\frac{1}{2}(x_3)^2$$

这些收益函数呈山形。求出一个纳什均衡。

12. 参与者 1,2 和 3 正进行一场博弈，其中参与者 i 的策略用 y_i 表示，y_i 可以是任一非负实数。参与者 1 的收益函数是：

$$V_1(y_1,y_2,y_3)=y_1+y_1y_2-(y_1)^2$$

参与者 2 的收益函数是：

$$V_2(y_1,y_2,y_3)=y_2+y_1y_2-(y_2)^2$$

参与者 3 的收益函数是：

$$V_3(y_1,y_2,y_3)=(10-y_1-y_2-y_3)y_3$$

这些收益函数呈山形，求出一个纳什均衡。（提示：参与者 1 和 2 的收益函数是对称的。）

第

7

章

让他们猜不透：随机策略

学习目标

7.1 在不确定下做决定

7.2 混合策略和纳什均衡

7.3 解决混合策略纳什均衡问题的提示

7.4 进阶提升

7.5 纯冲突和防范行为博弈

本章导读

“很明显，闯入者会在夜间行动，然后在白天出来，而罗伯特一直在他的小屋里保持清醒。问题是他是否应该修改计划：即给闯入者他白天睡觉而晚上保持清醒的印象？为什么呢？闯入者可以轻易改变他的策略。所以答案是不，罗伯特应该变得不可预测，以使闯入者变得不确定。当他清醒的时候应该假装睡着，而当他睡着的时候应假装清醒……他应该试着去考虑对方会认为他是怎么想的，或者是试想下对方会怎么考虑他是怎么考虑对方对他的考虑……”

为了扼制毒品交易，市长会增加巡逻警力来破坏毒贩的买卖。毒贩既可以在附近的街角从事毒品交易，也可以在附近的公园内从事毒品交易。每天，他都要决定在哪里设立贩毒点，他知道购买者会知道他所选择的地点是哪。因为没有一个很好的告密者，警察不会知道这个地点，所以警长也要决定自己是去公园还是去街角。

警长和毒贩的决策决定着当天毒品交易的程度。试想一下，如果没有警察的干预，将会有 100 宗毒品交易发生。一个毒贩所得的收益等于他所完成的交易数量。而对一位警长而言，他的收益是他所破获的交易数量（即用 100 减去实际发生的交易量）。如果他们都选择了公园，警长在那巡逻且毒贩在那活动，就只会发生 40 宗交易，也就是说 60 宗交易会被破坏。因为公园的范围较大，所以仍然有相当一部分交易会发生。如图 7—1 所示，此时警长的收益是 60，而毒贩的为 40。如果警长在街角巡逻时，毒贩在公园内活动，此时会有 100 宗交易发生，于是警长的收益为 0，而毒贩的收益为 100。如果警长在街角巡逻而毒贩碰巧在街角，那么只有 20 宗交易会发生。最后，如果毒贩在街角，警长在公园，会有 90 宗交易发生（因为毒贩的交易会被警长的巡逻车辆扰乱）。所有的收益情况如图 7—1 所示。

这个博弈中没有纳什均衡。例如，如果毒贩选择在公园，那么警长也

PDG

会选择去那。但若毒贩预测警长会在公园出现，他们会选择去街角。

		毒品交易	
		街角	公园
警车巡逻	街角	80,20	0,100
	公园	10,90	60,40

图 7—1　警车巡逻与毒品交易

这是一个以机智取胜的博弈例子，在第 4 章就曾介绍过。每个参与者都会选择一个对方无法预测的地点。毒贩不会总选择去公园，因为这种预测会导致警长去公园巡逻并破坏他们的交易。出于同样的原因，毒贩也同样不会总选择去街角。所以对毒贩来说，他们很自然的选择便是常常更换地点，有时去街角，有时去公园。但我们能准确地说他们将会如何更换地点吗？倘若毒贩经常更换他们的地点，警长又会怎么做呢？很显然，警长同样不希望自己被预测到，这是因为若警长总是去公园，那么毒贩就会非常轻松的在街角进行交易。

本章的目标是学习如何解决以机智取胜型博弈问题，或者更简单地说，从参与者使他们的行为随机化中来推理解决途径。为了实现这个目标，我们需要知道怎么建立一个在不确定下做决策的模型，这也是 7.1 小节中将分析的。有了这方面的知识后，我们将在 7.2 小节中回到毒品交易的问题并说明怎么解决。7.3 小节中给出了更多的例子，以及获得解决办法的更多提示。随后，更多的博弈问题在 7.4 小节中得以解决。在 7.5 小节中，具有随机参与性的解决途径的一种特性会在一些纯对立博弈问题中被推理出来。

7.1 在不确定下做决定

如果毒贩在选择去公园还是去街道贩卖时是随机决定的，那么警长在选择去哪巡逻时面临着一个不确定的结果。如果他在公园巡逻，那么他成功破坏毒品交易的程度取决于毒贩是否会在公园设立交易点。为了确定最优策略，警长会在不确定的结果下评估选择。想出一种在不确定下做决定的方法是本节的任务。

第一步是量化不确定性，这一步可通过 7.1.1 介绍的可能性和预期的概念来完成。这些概念又将会应用于 7.1.2 节的一种在不确定下比较不同选择的方法中。最后，在 7.1.3 中，我们将讨论当存在不确定性时，与理解收益（或效用）相关的重要问题。

7.1.1　概率和期望

在智能社区中可以用行话来描述不同事件的可能性。例如，国家会在未来 5 年里发展核武器是具有最小的可能性的事件，我们称之为“可信的”（conceivable）。如果认为这一事件更有可能性，但其发生的概率又低于 50%，我们描述这一事件为“也许”（possible）。如果此事件发生的可能性超过 50%，则被称之为“可能”（probable）。而当此概率在 60% ～ 70% 这个范围内时，我们称之为“很有可能”（likely）。当概率达到 80% 时，这一事件就叫做“基本确定”（almost certain）。当要表达最高可能性，例如当某事件一定会发生时，有人会说“毫无疑问”（no doubt）。

除了借助以上这些术语，我们往往需要更加准确的量化不确定性。为此，可能性概念的引出会非常有用。为学习到底什么是可能性，让我们先去拉斯维加斯游览一番，玩一下轮盘赌。

美国式轮盘上有 38 个数字：0，00 和 1 ～ 36。一个小球会在轮上旋转并最终停在 38 个孔中的一个内，代表一个不同的数字。[①]在一次特定旋转下被“击中”的数字被称为**随机事件**（random event），这也就意味着这是极不具可测性的。当然，原则上也有人会根据轮盘旋转的力度，旋转的速度以及球的起始位置进行预测。机灵的赌徒甚至还会引入更多的变量，例如房间的温度和湿度。但实际上，小球的落脚点远非人或机器所能控制的。

即便如此，我们虽不能预测球的最终位置，但我

①因为赌场只会为博中非零数字的赌客支付筹码，因此会有5.26%的概率为赌场带来利益。相对而言，欧洲式轮盘只有37个数字，这其中仅含一个数字零，因而其概率也只有2.7%。——作者注

们可以观察发生了什么。那么，我们来看看轮盘上的哪个数字是最容易被击中的。经过 1 000 次的试验，假设数字 17 被击中了 30 次。那么事件 17 被击中的**频率**（frequency）就是 30，且它的**相对频率**（relative frequency）——即相对于总次数的频率，就是30/1 000或者是0.03。在另一轮 1 000 次旋转中，我们发现 17 被击中了 24 次，即所得频率在第二轮 1 000 次旋转中为 0.024，那么前 2 000 次中的相对频率为 0.027。让我们继续旋转轮盘并记录结果。当旋转的次数趋于无限时，我们把 17 被击中的相对频率称为 17 发生的可能性。如果这个轮盘是“公平”的——每个数字被击中的可能性是一样的，那么经多次试验后 17 的相对频率将接近 1/38（因为轮盘上一共有 38 个数字）。

因此，**事件的可能性（probability）就是这个事件在无限次试验中发生的频率。**这样定义的话，可能性就是一个抽象的概念，因为我们无法进行一个无限次的试验来真正的衡量它。随机事件的可能性分配是随机事件所有可能结果的可能性集合。在轮盘赌的例子中，随机事件是一个特定的数字被击中，而可能出现的结果是 38 个数字，这种可能性分配就是 38 种可能性的一个集合。

一个有效的可能性分配需要满足两个条件：

（1）可能性为 0 或者 1，0 表示不可能发生，而 1 表示此事件肯定会发生。

（2）所有可能性的总和为 1。

想想轮盘赌，小球肯定会掉在某个数字孔中。因此，如果我们把所有数字的可能性加总，那么结果就是 1。如果不是那样的话，例如加总后为 0.98，那么我们就可以说有 2% 的可能性那个被击中的数字不是 0,00,1,2…,36。但这是没有道理的，因为我们已经规定了一个框架，而这个球总是会掉进这个框架中 38 个数字孔中的任意一个中。当然，你也许会说球可能会飞起来然后落在地上，这是一种物理上的可能性。但在这种情况下，我们又会把结果可能性拓展为既包括落在地上又包括落在 38 个数字孔中的一个中。一旦那种落在地上的结果被考虑进去，那么目前 39 种可能性的总和一定还是 1。换句话说，某事件一定发生，并且我们能具体衡量事件发生的可行模式，那么这一事件肯定是可能发生的 39 种事件中的一种。

当一事件发生的结果与另一事件发生的可能性没有联系时我们称这两个（或者更多）随机事件为**独立的**（independent）。让我们考虑这样一个例子。假设我们有两个轮盘，一个是美国式的，一个是欧洲式的，它们都是公平的。我们定义两个事件：在美国式轮盘被击中的数字和在欧洲式轮盘被击中的数字。那么在美国式轮盘击中数字 9 且同时在欧洲式轮盘击中数字 23 的可能性是多少呢？因为在美国式轮盘上所得到的结果不会影响在欧洲式轮盘所得到的结果，所以这两事件是独立的，并且欧洲式轮盘上被击中的数字的可能性与不考虑美国式轮盘上被击中的数字时是一样的。

当两事件独立时，两事件共同发生的可能是就是两事件分别发生的可能性的乘积。例如，在美国式轮盘上得到 9 和同时在欧洲式轮盘上得到 23 的可能性就等同于在美国式轮盘上得到 9 的可能性乘上在欧洲式轮盘上得到 23 的可能性，即（1/38）×（1/37）。

但事件又并非总是独立的。设想这样一个例子，两个人要进行一次测试，某人的得分是一次随机事件。如果他们是陌生人，他们所得分数的事件很有可能是独立的，即一个人的得分与另外一个人的得分没有任何联系。但假若他们是朋友或者在一起学习，并且选择同一份学习资料。在这种情况下，他们的得分的事件就不是独立的。如果他们其中的一个得分更高，说明他选对了学习资料。而另一个人选择同样的学习资料，他就也很有可能得到一个很高的分数。

随机变量（random variable）是随机事件取得的数值。明天的高温是一个随机变量，这和道琼斯指数的数值很接近。随机变量的**期望值**（expected value，或者叫期望，expetation）是随机变量可能实现值的加权，而事件可实现性的加权即事件发生的可能性。例如，假设你在轮盘赌桌上在数字 14 上押了 10 美元，而在数字 32 上押了 5 美元。利益的随机变量是你所赢得的总量。比如说在一个美国式轮盘赌上投注 36 美元，每一美元投在一个不同的数字上，所赢总量的期望为：

$$\left(\frac{1}{38}\right)\times 10\times 36+\left(\frac{1}{38}\right)\times 5\times 36+\left(\frac{36}{38}\right)\times 0$$

$$=\left(\frac{1}{38}\right)\times 360+\left(\frac{1}{38}\right)\times 180=\frac{540}{38}\approx 14.21$$

数字 14 被击中的概率是 1/38，这时你将赢得 360 美元（包括下注的 10 美元）。数字 32 有 1/38 的概率被击中，此时你赢得 180 美元。在剩下的 36/38 的概率中，数字既不是 14 又不是 32，你的所得为 0。因此，你所

赢总量的期望为 540/38。

基于概率的定义，期望值可以看成是事件经过无数次试验后的一种平均值。如果你再一次下赌注，你会发现最终你将一无所得，既没有 180 美元，也没有 360 美元。如果你在轮盘上多次下注，你的平均胜率是 540/38，或者是 14.21 美元，这将低于你付出的 15 美元的平均赌注，这样你就注定要输钱，这也就是赌场挣钱的道理。

7.1.2 不确定选择中的偏好

星期六晚上，黛安娜要决定去哪吃晚餐。她可以去她经常去的地方——福瑞欧烧烤店，一家食物不错，价格也适中的店。她去过那很多次了，也知道那的具体情况。或者她还可以去尝试下新开的、高档的普罗旺斯什锦菜饭店。这是一家在当地享有良好口碑的饭店，但她不确定这家饭店是不是名副其实，所以这可能是一顿价格过高的晚餐，也可能是一顿令人难忘的晚餐。

表 7—1 列出了所有黛安娜可能体验的餐馆的不同的收益（或者效用）。（目前，可以忽略"可选择性收益"一栏。）福瑞欧烧烤店的收益率是 100，戴安娜已经去过那很多次了，这是一个确定的事实。如果普罗旺斯什锦菜饭店能达到如外界评论的效果，则收益是 180。如果普罗旺斯什锦菜饭店实际上是价格过高且虚有其表的，则收益为 60。根据以往外界对事物评论的经验，黛安娜相信普罗旺斯什锦菜饭店有 60% 的可能性能够提供一桌美妙的事物。

表 7—1 黛安娜的餐厅选择

餐厅	晚餐质量	可能性	原始收益	可选择性收益
福瑞欧烧烤店	正常水平	1.0	100	100
普罗旺斯什锦菜饭店	价格过高且虚有其表	0.4	60	60
普罗旺斯什锦菜饭店	令人赞叹	0.6	180	120

所以，黛安娜该如何选择呢？为了回答这个问题，一个关于如何在不确定下做决策的理论是必不可少的。如果现在面临 3 种确定的选择：在福瑞欧烧烤店吃晚餐，或是在普罗旺斯什锦菜饭店吃一顿不错的晚餐，亦或在普罗旺斯什锦菜饭店享受一顿美妙的晚餐——那黛安娜的选择应该是第三种，因为这将给黛安娜最高的收益：180。但这些都不是她的选择。她的选择是要么去福瑞欧烧烤店获得一顿确定有保证的晚餐，要么冒着风险去普罗旺斯什锦菜饭店，要么两家都不去。

回顾第 1 章，我们得知当一个人的偏好得以实现后，她的行为可以被描述为一种选择最高收益（或效用）的行为。在不确定条件下，存在一种相关的原理：**对于一组特定的条件（这组条件不言自明），一个人的选择可以被描述为选择最高的期望收益的行为。**由于收益是个随机变量，期望收益可以用 7.1.1 小节所列举的公式进行计算。去福瑞欧烧烤店没有任何不确定性，所以从那取得的收益是 100。对于普罗旺斯什锦菜饭店而言，则是 $0.4\times60+0.6\times180=132$。因此，根据期望收益准则，黛安娜会选择去普罗旺斯。

于是，我们可以这样认为，当面临着包含随机结果的不同选择时，决策者会选择带来最高期望收益的那个选择。

7.1.3 序数效用对基数效用

在 2.5 小节所探讨的情形中，决策者的选择只取决于他对不同选择等级的高低排序。例如，回想托斯卡的情况（2.4 节），如图 7—2 所重现。当她刺伤斯卡比亚而斯卡比亚用空弹时（让她的情人卡瓦拉多西活着），托斯卡的效用为 4。当她屈服于斯卡比亚，而斯托比亚用空弹时，托斯卡的效用为 3。当她刺伤斯卡比亚，而斯卡比亚用真实子弹射杀她的情人卡瓦拉多西时，托斯卡的效用为 2。当她屈服于斯卡比亚，而斯卡比亚又使用真子弹时，托斯卡的效用为 1。在第 2 章，我们的结论是托卡斯应刺杀斯卡比亚。若斯卡比亚选择空弹，则托斯卡选择刺杀的效用为 4，选择屈服的效用为 3；如果斯卡比亚选择真实子弹，那么，托斯卡还是选择刺杀更好，因为效用为 2，而选择屈服的效用为 1。

		斯卡比亚	
		真子弹	空弹
托斯卡	刺杀	2,2	4,1
	屈服	1,4	3,3

图 7—2 托斯卡

目前，要重点关注的是只要我们保持结果的排序不变，刺杀是托斯卡的最优选择，这不会随着某效用大小的改变而改变。举个例子来说，结果（屈服，空弹）所得效用从 3 变为 3.99，托斯卡仍然会认为最好的方案是刺杀斯卡比亚；或者是若组合（刺杀，真子弹）的效用

从 2 变为 1.001，刺杀仍然是一个主导性的决策。

当效用仅包含一个人如何对不同选择进行排列这一个信息时，这种效用才是有序的。特别的，效用自身的量值是没有意义的，即说 3.99 比 3 更接近于 4 是没有意义的。这些都叫做序数效用，仅需注意 3.99 小于 4（正如 3 小于 4 一样）即可。同样，这也就是托斯卡更倾向于选择组合（刺杀，空弹），而非（屈服，空弹）的原因。为了得出托斯卡所应采取的合理选择，我们只需知道她对所有结果正确排序即可。

当选择中包含不确定结果时，再靠决策者对不同结果的排序来判断就不行了。要弄清其中原因，我们最好引出一个例子。回到图 7—1，考虑黛安娜不同的可选择性效用，和原始效用不同的一栏是去普罗旺斯什锦菜饭店享受令人赞叹的晚餐的可选择性效用，此时的可选择性效用为 120，低于原始效用 180。这两种效用都位于同一等级。当黛安娜选择去普罗旺斯什锦菜饭店享受晚餐时，不管效用为 120 或是 180，去那都是她的第一选择。现在我们考虑当存在有普罗旺斯什锦菜饭店食物不确定性的情况下，黛安娜对于福瑞欧烧烤店和普罗旺斯什锦菜饭店的选择。从原始效用看，从普罗旺斯什锦菜饭店那获得的期望效用为 132，所以黛安娜选择去那。当有选择性效用存在时，期望效用就变为 $0.4\times60+0.6\times120=96$，低于 100，她自然会选择福瑞欧烧烤店。普罗旺斯什锦菜饭店丰盛大餐效用的改变改变了黛安娜的选择，即便是在我们保持排序不变的条件下。

从上面的例子中我们可以得出：当面临不确定因素时，效用等级开始起作用，而不仅仅是效用所暗含的结果排序。效用不仅暗示着排序，也包含了偏好的强烈程度。当把黛安娜的效用从 180 降低至 120 时，也同时减弱了她对于想去普罗旺斯什锦菜饭店享受晚餐的欲望。在效用为 120 时，比起福瑞欧烧烤店，去普罗旺斯什锦菜饭店的驱动力并没有那么强烈。当效用的相应等级包含信息时，这时称效用是基数的。

当我们试图将某个参与者的序数效用合情合理的排列出来时，我们需要依靠对参与者实际情况的了解来排列他得到的权益。但对于基数效用，特别是不确定的情况下的基数效用，我们的任务是尽可能多的考虑需求。假如要恰当地描述一个人偏好的强烈程度，就需要我们赋予其各个效用一个相对的值。①

7.2 混合策略和纳什均衡

当我们掌握了在不确定情况下做决策的方法时，我们将在 7.2.1 小节里德警长和毒贩的对峙中介绍一种方法使参与者随机化。这一解决方法的一般特性将会在 7.2.2 小节中进行介绍。

7.2.1 回到开始？

有了期望收益的准则后，我们可以返回去探寻警长和毒贩会怎么做。图 7—1 所示的效用现在可以理解为是基数的而且能够反映出偏好的强烈程度。因为可以认为毒贩的效用等于毒贩所做成的交易数，所以更大数量的交易意味着会产生一个更加诱人的结果。同样的，对于警长来说，他也会认为更多数量代表着更大价值，因为她所获得的效用是她所破获的交易的数量。

为了在允许参与者使他们的行为随机化的情况下寻求解决办法，第一步就是重新定义这个游戏规则以允许这种可能性的存在。回想下，每个参与者都有两种决策：街道和公园。我们把这些叫做纯策略。每位参与者有使他的纯策略随机化的选择，我们称之为混合策略。具体来说，在这个博弈中，混合策略是区间 [0,1] 的一个实数，这个数值也就是选择街道的可能性。对警长而言，用 p 表示她去街道巡逻的概率，$1-p$ 表示她选择去公园的概率。类似的，毒贩的混合策略中用 d 表示他选择去街道的概率，$1-d$ 表示他选择去公园的概率。

为了完善这些新策略组合，我们需要重新定义这些决策的衡量方法。首先，我们用期望收益替换表 7—1 中的原始收益。为推导出期望收益，我们写下不同结果下的概率，如策略组合(p,d)。这些概率如表格 7—2 所示。因为警长和毒贩的随机策略是独立的，也就是说警长和毒贩都选择去公园的概率等于警长去公园的概率乘以毒贩去公园的概率。

①任何相关联的效用变换，不会改变期望效用标准及其作用下的参与者的决策行为。仿射变换（affine transformation），就是以某个常量分别累加于各个效用，或是以某个正数分别乘以各个效用。举个例子，假如将戴安娜的三个效用分别加上63，其在不确定性下的饭店选择不会随之改变。再如，将这3个效用分别乘以71，她的选择还是不会改变。事实上，这种坐标的变换不会导致结果的改变。如果效用是用美元计量的，那么换做便士（即效用乘以100倍），也不会影响到某人最终的选择。效用坐标系的变换就像是温度的变换，不论是以华氏度计量还是以摄氏度计量，我们感知到的冷暖不变。这和效用是同一回事。——作者注

表 7—2　警车巡逻与毒贩交易博弈的结果

警长的选择	毒贩的选择	可能性	警长的收益	毒贩的收益
街道	街道	$p\times d$	80	20
街道	公园	$p\times(1-d)$	0	100
公园	街道	$(1-p)\times d$	10	90
公园	公园	$(1-p)\times(1-d)$	60	40

策略组合（p,d），即警长的期望收益，用 $V_{PO}(p,q)$ 表示就是：

$$V_{PO}(p,d)=p\times d\times 80+p\times(1-d)\times 0+(1-p)\times d\times 10+(1-p)\times(1-d)\times 60$$
$$=60-60p-50d+130pd$$

例如，概率 $p\times d$ 表示警长和毒贩都去街道上，在这种情况下，警长的效用为 80，而毒贩的期望收益为：

$$V_{DD}(p,d)=p\times d\times 20+p\times(1-d)\times 100+(1-p)\times d\times 90+(1-p)\times(1-d)\times 40$$
$$=40+60p+50d-130pd$$

为得出结论，我们用两种方法改变了图 7—1 中的博弈规则。

> 第一，决策由两种纯策略变为一种随机化策略——我们称之为混合策略，且这一策略组合是用不同方法通过这些纯策略来分配概率的。
>
> 第二，混合策略的估值根据它所产生的期望效用来实现。

被改造后的博弈被称为博弈的随机版本。我们从这一改造中学到了什么呢？答案是尽管图 7—1 中的最初博弈中没有纳什均衡，但在改造后的随机版本中却出现了纳什均衡。

为了从随机版本中推导出解决办法，我们首先从警长的两个纯策略中找出期望效用，并说明这两个纯策略在值 d 上的独立性。值得注意的是一个纯策略仅仅是混合策略的一种特殊（或退化）情况。因此，纯策略选择街道相当于条件 p=1，纯策略选择公园相当于条件 p=0。在这种情况下，有：

$$V_{PO}(1,d)=60-60-50d+130d=80d$$

图 7—3 中给出了这一期望效用以及选择公园的期望效用，即：

$$V_{PO}(0,d)=60-50d$$

在图 7—3 中，我们可以推导出警长的最佳策略函数，可以描述她在每个值 d 下的最佳值 p。从图 7—3 中很容易看出，当 $d<6/13$ 时，她倾向于去公园巡逻（即：$p=0$ 的期望效用超过 $p=1$ 的期望效用）。换句话说，如果毒贩非常不可能去街道上贩毒时（即 $d<6/13$），警长会倾向于去公园巡逻。

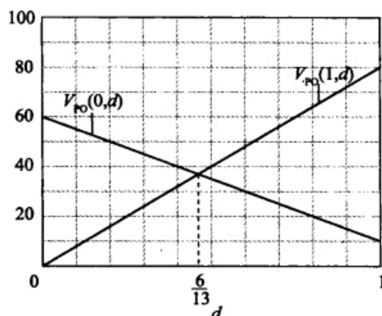


图 7—3　警长的期望效用

至此为止，我们已经说明如果 $d<6/13$，则 $p=0$ 时的选择优于 $p=1$ 时的选择。接下来，我们继续说明如果 $d<6/13$，则 $p=0$ 条件下的选择仍然优于 p 为其他任何值的选择，即：

$$V_{PO}(0,d)>V_{PO}(p,d),p>0 \quad (7.1)$$

我们重新整理了警长的期望效用，得到：

$$V_{PO}(p,d)=p\times d\times 80+p\times(1-d)\times 0+(1-p)\times d\times 10+(1-p)\times(1-d)\times 60$$
$$=p\times[d\times 80+(1-d)\times 0]+(1-p)\times[d\times 10+(1-d)\times 60]$$
$$=p\times V_{PO}(1,d)+(1-p)\times V_{PO}(0,d) \quad (7.2)$$

最后一行等式表示警长的期望效用来自策略 p，也就相当于巡逻街道的概率乘上出现在街道上的期望效用 $V_{PO}(1,d)$，再加上巡逻公园的概率乘上她这样做时在 $V_{PO}(0,d)$ 处的期望效用。由图 7—3 可知 $d<6/13$ 时，

$$V_{PO}(0,d)>V_{PO}(1,d) \quad (7.3)$$

现在我们在公式（7.3）上做一系列的代数变换，得到：

$$V_{PO}(0,d)>V_{PO}(1,d),$$
$$p\times V_{PO}(0,d)>p\times V_{PO}(1,d),$$
$$V_{PO}(0,d)>(1-p)\times V_{PO}(0,d)+p\times V_{PO}(1,d),$$
$$V_{PO}(0,d)>V_{PO}(p,d)$$

最后一行即公式（7.1），也就是我们想要证明的。

对于结果，我们并不需要感到很吃惊。因为 $V_{PO}(p,d)$ 是 $V_{PO}(1,d)$ 和 $V_{PO}(0,d)$ 的加权平均数，如果 $V_{PO}(0,d)$ 大于 $V_{PO}(1,d)$，那么 $V_{PO}(0,d)$ 就大于 $V_{PO}(1,d)$ 和 $V_{PO}(0,d)$ 的加权平均数。总而言之，如果 $d<6/13$，那么 $p=0$ 就会产生比 $p>0$ 的期望效用更大的期望效用，因此 $p=0$ 是警长的唯一最佳选择。

用同样的逻辑，我们可以得出如果值 $d>6/13$，那么 $p=1$（确定选择街道）不仅会产生出比 $p=0$ 时更高的期望效用，而且也会产生比 $p<1$ 时更高的期望效用，因此这将是警长的最佳策略。用直觉来看，如果毒贩更有可能去街道，那警长也更愿意去那里。

最后留下的情况是当 $d=6/13$ 时。在这种情况下，选择纯策略街道和纯策略公园所得的期望效用是一样的。进一步说，无论 p 为任何值，期望效用都是一样的。这也是因为：

$$V_{PO}\left(p,\frac{6}{13}\right)=p\times V_{PO}\left(1,\frac{6}{13}\right)+(1-p)\times V_{PO}\left(0,\frac{6}{13}\right)$$
$$=V_{PO}\left(0,\frac{6}{13}\right)+p\times\left[V_{PO}\left(1,\frac{6}{13}\right)-V_{PO}\left(0,\frac{6}{13}\right)\right]$$
$$=V_{PO}\left(0,\frac{6}{13}\right)$$

前面提到 $V_{PO}\left(1,\frac{6}{13}\right)=V_{PO}\left(0,\frac{6}{13}\right)$，所以表 7—2 第二行的第二个值是 0。因为对于不同的 p 值 $V_{PO}\left(p,\frac{6}{13}\right)$ 是不变的，所以属于区间 [0,1] 的 p 都是一种最佳策略。

把所得结论放到一起，我们在图 7—4 中得出警长最佳策略的曲线。在 $d=6/13$ 处从 0 到 1 的横线表明在这种情况下所有的 p 值都是最佳选择。

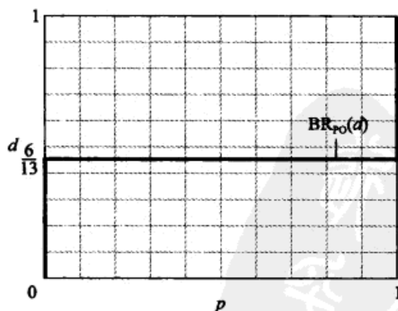


图 7—4　警长最佳策略

同理毒贩的最佳策略可以从图 7—5 中得到，图 7—5 中的另一条曲线是警长的最佳策略选择曲线。（作为练习，请自己完成）。当 $p<5/13$ 时，在街道上贩毒的纯策略对于毒贩来说是最佳的选择。如果警长极有可能不会去街道巡逻，那么这就是毒贩最想去的地方。如果 $p>5/13$，那么出现在公园（$d=0$）的纯策略将是唯一的最佳策略。而当 $p=5/13$ 时，毒贩对于两种纯策略的选择是无所谓的，这就意味着此时任何 d 值都是最佳策略。

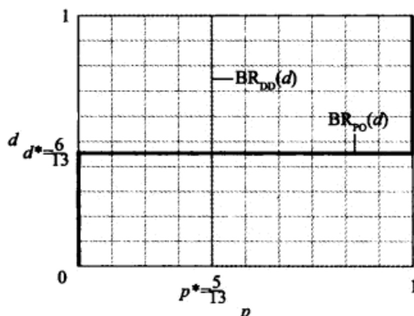


图 7—5　警车巡逻和毒贩交易的纳什均衡

前面说过，当参与者的策略都是相对于对方策略的最佳策略时，我们称之为纳什均衡的策略。用 $BR_{PO}(d)$ 和 $BR_{DD}(p)$ 表示警长和毒贩的最佳策略，相应的，我们可以看出如果有 $p^*=BR_{PO}(d^*)$ 和 $d^*=BR_{DD}(p^*)$ 就是纳什均衡。

这是两个最佳策略函数的交点，在图 7—5 中已经描绘出来。倘若毒贩选择混合策略 d^*，则混合策略 p^* 就是警长的最佳策略。进一步，倘若警长选择最佳策略 p^*，则混合策略 d^* 就是毒贩的最佳策略。从图 7—5 中可以清楚看到纳什均衡，它表明警长在街道巡逻的时间占 $\frac{5}{13}$，毒贩在街道的时间占 $\frac{6}{13}$。

7.1 理解测验

在图 7—6 中找到混合策略中的纳什均衡。

		参与者2	
		左	右
参与者1	顶部	8,2	1,6
	底部	3,5	4,1

图 7—6

7.2.2　纳什均衡在混合策略中的一般属性

让我们回顾下刚学的知识。在已给出的一个博弈中，用 m_i 表示参与者 i 的纯策略数量。在一个随机化博弈

类型中，一策略应为参与者所有纯策略中的随机化策略。因此，将参与者 i 的一个混合策略的概率分配到他的每一个纯策略 m_i 中。混合策略集合中一组 m_i 的概率属于区间 [0,1]，其所有概率总和为 1。换句话说，这一混合策略集合包含所有的参与者纯策略的随机化。值得注意的是混合策略组也包含了纯策略。每位参与者用相关的期望收益来评估一特定的混合策略。随机化博弈中的纳什均衡就是一个混合策略，每位参与者的混合策略是他们的期望收益最大化，同时也赋予其他参与者混合策略。

在进行一些具体应用之前——这是做这个分析最有趣的事，要注意混合策略纳什均衡中的一些关键点。首先要感谢约翰・纳什，这个我们在第 4 章已有提及。回想下一个有限的博弈是由一群有限的参与者进行的博弈，对于每位参与者而言，都是一些有限的（纯）策略。


点拨

纳什均衡的存在：每个有限的博弈的混合策略中都有纳什均衡。

这是一条非常有用的性质。因为它告诉我们如果要进行一次正确的分析，在混合策略中找出纳什均衡尤为重要，而我们总能找到这一均衡。这一均衡也许是纯策略中的一个等式，或者是至少一名参与者的随机量。

在建立了至少存在一个纳什均衡的博弈体系后，我们关注的另外一点是斯坦福大学的罗伯特・威尔森（Robert Wilson）提出的均衡数量论，即到底存在多少个这样的均衡等式。


点拨

纳什均衡的数量：在几乎所有的有限博弈中，总有有限个、奇数个纳什均衡等式存在于混合策略中。

除了一些特殊的博弈外，在大多数博弈中，纳什均衡等式的数量都是奇数。这件事听起来玄乎其玄、非常奇怪，但非常有用。除非你研究的博弈非常特别，如果你发现比如说两个均衡等式，你就会知道至少还有一个以上的等式隐含其中。例如，我们在 4.2 节中求解常规博弈时发现两个纯策略纳什均衡等式。威尔森的理论表明也许还有更多的均衡等式，事实上，在特尔玛和路易斯随机化中还存在第三个纳什均衡。

7.3 解决混合策略纳什均衡问题的提示

这节将会列出一些例子来帮助你更好地理解混合策略纳什均衡（当然这并不是一件容易的事），然后展现这些混合策略的相关性以为不同策略背景中的具体应用提供解决途径。我们先给出一些解决混合策略纳什均衡问题的简单提示。

在毒品交易博弈中，毒贩在街道的概率是 $\frac{6}{13}$，在公园的概率为 $\frac{7}{13}$ 是一个均衡。毒贩的随机化是比较合理的，因为这将使警察无从入手。从另一方面来说，你会好奇毒贩怎样选择他们的策略才能使随机策略最优化，即怎么用一枚硬币的不同翻转来决定你应该怎么做才能使随机策略最优化。

尽管通常来说，用投硬币的方法来决定自己的行为并不是最好的选择。但这里存在一个期望：即当你对你面临的选择都漠不关心的时候，随机化一个选择集合无疑是一个最佳选择。如果在街道上或者是在公园里都有着同样的期望收益，那么毒贩就不会关心其作出的决策，也会对决定他命运的随机策略非常满意。这就是参与者选择随机化时出现纳什均衡的例子。当警察在街角巡逻的概率为 $\frac{5}{13}$，毒贩任意选择的纯策略的期望收益都为 $\frac{820}{13}$。

毒贩选择街道的期望收益：

$$\left(\frac{5}{13}\right)\times 20+\left(\frac{8}{13}\right)\times 90=\frac{820}{13}$$

毒贩选择公园的期望收益：

$$\left(\frac{5}{13}\right)\times 100+\left(\frac{8}{13}\right)\times 40=\frac{820}{13}$$

这种情况下，毒贩选择在街角的概率为 $\frac{6}{13}$，对他来说无疑是最佳选择（就如同其他任何混合策略一样）。


点拨

如果在纯策略集中使其随机化是最优选择，那么参与者必须从任意纯策略组中获得相同的期望收益。

这一理解为我们解决混合策略纳什均衡提供了一个

非常有用的方法。如果毒贩选择随机化其策略，那么警察必须选择混合策略才能使毒贩作出任意选择的机会均等。那就是说，警察的均衡策略的赋值 p，与毒贩选择的两个纯策略的期望收益相等：

$$p\times20+(1-p)\times90=p\times100+(1-p)\times40\Rightarrow p=\frac{5}{13}$$

然后找出毒贩的均衡策略，你需要推导出使警察的两个纯策略机会均等的 d 的值：

$$d\times80+(1-d)\times0=d\times10+(1-d)\times60\Rightarrow d=\frac{6}{13}$$

诚然我们用警察的收益去决定毒贩的均衡策略，而用毒贩的收益来决定警察的均衡策略似乎有些奇怪。但我们必须牢记纳什均衡在混合策略里的特性。一个参与者只有在策略会被随机化的机会均等的情况下才会最优的选择随机化。因此，这种方法需要为参与者找到一种策略使得其他参与者认为自己的纯策略（或纯策略的某些集合）机会均等。

因为解决纳什均衡包含许多复杂的代数应用，任何能使这个问题简单化的途径都是非常有用的。这里有一些适用于这个问题的技巧：**如果一个参与者有一个严格劣势（纯）策略，那么此策略在纳什均衡混合策略中被赋予的概率值为0。**我们要记住的是，参与者永远不会按照随机机制暗示的那样去执行。因此，要想使选择达到随机机制规定的最优化，此纯策略必定在所有纯策略中达到最高的收益，因此当策略为严格劣势策略时，这一目标是无法实现的。

为了试图解决混合策略纳什均衡问题，我们需要剔除所有的严格劣势策略。剔除后的博弈中的混合策略纳什均衡组将会和最原始的博弈一模一样，因为我们剔除掉了所有不能被用于应用中的纯策略（也就是说它们的概率为0）。同理，如果在一个剔除后的博弈中存在一个严格劣势策略，一个纳什均衡必须赋予它0概率，否则这些策略将会被其他参与者使用，严格劣势策略是劣于一些纯策略的。在原始策略中剔除所有严格劣势策略的博弈中，纳什均衡策略将赋予这些严格劣势策略0概率。这些论断使得我们可以得出一个更加精简的博弈，这个博弈使得我们可以推断出严格劣势策略应用其中时，其被赋予的概率为0。顺着这个逻辑，我们可以得出当一策略无法在重复剔除严格劣势策略后继续使用时，纳什均衡策略赋予其的概率为0。


点拨

若一纯策略无法在重复剔除严格劣势策略后继续使用时，纳什均衡策略赋予其的概率为0。

我们的建议是首先重复剔除严格劣势策略来排除所有被赋予0概率的策略，然后从余下的博弈策略中找出混合策略纳什均衡。

案例研究

第二次世界大战中的阿维让切斯防线

1944年的诺曼底登陆是20世纪军事史中最著名的事件。许多书中以及诸如斯皮尔伯格的《拯救大兵瑞恩》（*Saving Private Ryan*）等许多影视作品中，都曾描写过当时用以击败德国纳粹的手段和策略。在我们所熟知的代号为“操控上帝”的行动中，策略是关键因素，特别值得注意的还有它的日期和地点。德国人预期到会有一场入侵，但问题是什么时候和在哪儿？联军知道使用双面特工和无线电通信时，信息会被德国人截获，所以联军下令欺骗德国人，让他们以为联军会在法国和挪威发起入侵行动。

这种策略不仅会在这样的高级行动中取得良好效果，也会在一些普通行动中生效。许多军官将领在战争开始时就面临这些独立的选择策略的情况，其中的一个例子就出现在阿维让切斯防线事件中。阿维让切斯是法国海岸线上的一处区域，当时，欧马·布拉德利（Omar Bradley）将军指挥美国第一军队，与德军军官冯·克鲁格（Von Kluge）指挥的德国第七军队在阿维让切斯对决。值得我们关注的是，布拉德利在防线拥有四个分队的情况下到底应该怎么做。布拉德利认为他面临两种选择：首先，他可以下令军队加固防线，使得德军想要攻破防线更加困难。其次，他可以使得防线向东移，其目的是为了累垮德军。用他自己的话来说就是：

> 我们要么召回余下的四个分队来加强芒什省的霍奇斯防线来保卫英军的生命战线，这是用一种安全的方法来加固防线；要么通过命令军队突破霍奇斯防线与暴露在外的德军士兵殊

死一搏。



对于德国的冯·克鲁格将军来说，我们不知道他到底在想什么，但他的两个合理选择是进攻或者撤退。以下是布拉德利当时描述的：

> 德国军官面临一个很困难的决定，他要么撤出余部从而露出两翼，加强他的南北线，这样可以让他的部队完整有序地撤到塞纳河；要么他放手一搏，让他的军队进攻并占领防线从而控制通向大海的防线。

图 7—7 展示了当时的地理情况，以及可能出现的四种可能结果。

图 7—8 表示策略设置的效用矩阵。对于布拉德利而言最好的结果是，克鲁格撤军，而他向东增派军队来拖垮德军。而对他来说最差的结果是（这时就是对克鲁格来说最好的结果），他向东增派军队，而克鲁格发动进攻，因为布拉德利的军队此时就很难抵御敌人强力的进攻。对克鲁格来说最差的结果是当布拉德利选择驻守防御时，他主动发起进攻，此时克鲁格的失败似乎是注定的。



图7—7　阿维让切防线可能的部队部署

		克鲁格	
		攻击	撤退
布拉德利	加固防线	3,0	2,3
	向东进军	0,5	4,2

图 7—8　阿维让切斯防线的情况

不难注意的是在这个博弈的纯策略中不存在纳什均衡。如果克鲁格选择进攻，那么布拉德利最佳的选择是防御。但如果布拉德利选择防御，克鲁格最佳的选择是撤退。反而言之，如果克鲁格选择撤退，布拉德利则很可能向东进军，但如果是这样的话，克鲁格应该选择进攻，而不是撤退。在不知道任何一名军官的选择之前，都不存在稳定的纯策略最优选择。

为了在混合策略中找出纳什均衡，用 r 表示布拉德利选择驻军防守的概率，用 $1-r$ 表示他选择向东进军的概率。对克鲁格而言，用 a 表示他选择进攻的概率，用 $1-a$ 表示他选择撤退的概率。如果克鲁格是随机的选择，那么他对于两个选择中的任何一个都是机会均等的选择，即由以下式子表示：

$$r\times 0+(1-r)\times 5=r\times 3+(1-r)\times 2$$

等式坐标表示选择进攻的期望收益，等式右边表示选择撤退的期望收益。为了研究这个问题，假设 $r=1/2$。如果布拉德利选择向东进军和选择驻军防御的概率均等的话，那么不管克鲁格是选择进攻还是撤退，其期望收益是相同的。

如果布拉德利要想使他的随机化选择最理想，那么或是选择 $\frac{1}{2}$ 概率的向东进军或是其他概率，他认为对这两个纯策略的选择机会均等，即有：

$$a\times 3+(1-a)\times 2=a\times 0+(1-a)\times 4\Rightarrow a=\frac{2}{5}$$

因此，如果克鲁格进攻的概率为 $\frac{2}{5}$，那么布拉德利的对于他是否选择向东进军的期望收益都是相同的。那么，$a=\frac{2}{5}$，$r=\frac{1}{2}$ 就是最佳的选择（就和布拉德利选择其他任何混合策略是一样的）。进一步来说，$a=\frac{2}{5}$ 是对于 $r=\frac{1}{2}$ 时来说最佳的选择。因为我们已经知道，当布拉德利很有可能选择向东进军或不向东进军的时候，克鲁格此时对于他所做的随机化选择都是满意的。总之，$r=\frac{1}{2}$，$a=\frac{2}{5}$ 是一个纳什均衡，所以两位军官的选择都是随机化的。

那么，余下的故事该怎样继续呢？当然，它比我们分析所得出的启示更为复杂（因为这里明显存在第三参与者）。克鲁格选择撤退，但他收到阿道夫·希特勒的限制，即希特勒会坚持让克鲁格进攻。克鲁格执行了命令，进攻失败。撤退途中，他最终选择了自杀。但故事还没有结束。最近的发现揭露出布拉德利很早已经知道克鲁格所做的决定，因为联军有能力截获和解码德军秘密信息。

关于阿维让切斯防线更为合理的分析需要一个更为复杂和精准的模型。

进入市场

设想一个新市场诞生，四家有前景的公司在考虑着进入这个市场。进入的成本因公司的不同而不同，如表 7—3 所示。一旦他们进入，所有公司会取得相同的毛利润，而这利润只取决于总共有多少竞争者。这些公司的毛利润在表 7—4 列出。公司选择进入市场的收益是所得的毛利润减去进入的成本，不进入的收益为 0。

表 7—3　进入市场的不同成本

公司	进入成本
1	100
2	300
3	300
4	500

表 7—4　毛利润

进入者数量	每个进入者的毛利润
1	1000
2	400
3	250
4	150

在从此例中得出混合策略纳什均衡之前，我们需要确定这个博弈是否能被简化。一种办法是检查严格劣势策略，因为我们已经知道这些策略在纳什均衡中被赋予的概率为 0。对公司 1 来说，不进入严格劣于进入：不管有多少公司会选择进入市场，公司 1 的所得收益都是正的，因为它的毛利润不会低于 150，而它的进入成本只有 100。所以，此时纳什均衡赋予公司 1 不进入的概率为 0。那就意味着公司 1 会选择使用纯策略进入。换句话说，进入是公司 1 的占优策略。

接下来，注意到只要公司 1 选择进入，不管公司 2 或 3 无论怎么选择，公司 4 选择进入市场的收益都是负的。即便在公司 2 和 3 都不选择进入时，公司 4 的收益也为−100（=400−500），这就比选择不进入时的收益少。在纳什均衡作用下，公司 4 会选择不进入的纯策略。

在进行了两个回合重复剔除严格劣势策略后，我们可以总结出，在纳什均衡下，公司1会选择进入，而公司4会选择不进入。现在问题只剩下公司2和3的抉择了，见图7—9。

		公司3	
		进入	不进入
公司2	进入	-50,-50	100,0
	不进入	0,100	0,0

图7—9　主导策略分析排除后的进入博弈

这个博弈中有两个纯策略均衡：（进入，不进入），（不进入，进入）。总结来说就是，四人博弈仅有两个纯策略纳什均衡：（1）公司1和2进入，公司3和4不进入；（2）公司1和3进入，公司2和4不进入。

回顾之前所学，在大多数有限博弈论中，总有一确定的奇数数量的纳什均衡。到目前为止，我们找出两个均衡；但应该有更多均衡。此外，因为已经找出所有纯策略均衡，任何余下的纳什均衡必须包含至少一个公司是随机化的组成。

回到表7—9，假设公司2和3会随机化它们的进入决策，用e_i表示公司i（=2,3）选择进入市场的概率。那么使公司2均等选择进入或不进入的e_3被定义如下：

$$e_3\times(-50)+(1-e_3)\times100=0 \qquad (7.4)$$

式子左边表示选择进入的期望收益，式子右边0表示不进入的期望收益。如果公司3在进入概率为e_3的条件下选择进入，再加上公司2选择进入，此时就有3家公司在这一行业（主要注意的是公司1同样会选择进入），而此时公司2的期望收益为-50。在概率为$1-e_3$的情况下，公司3不会选择进入，而此时公司2选择进入的期望收益就为100。解出此等式，我们得到$e_3=\frac{2}{3}$，同样，公司2和3面临的情况是对称的，同样的推理可以得出如果$e_2=\frac{2}{3}$，公司会均等地去选择两种策略。

我们得到了第三个纳什均衡，即公司1肯定会选择进入，公司2和3选择进入的概率都为$\frac{2}{3}$，公司4肯定不会进入。由此，我们可以得出纳什均衡可以有一部分参与者的选择是随机化的，另一部分是固定的。

7.2 理解测验

在图7—10中找到混合策略中的纳什均衡。

		参与者2			
		w	*x*	*y*	*z*
参与者1	*a*	1,5	2,2	0,4	3,1
	b	3,0	4,1	2,5	4,2
	c	1,3	2,6	5,2	2,0

图7—10

7.4 进阶提升

以上的例子都是为了解决诸如有两个参与者且每位参与者有两种策略的混合策略。在这一节，我们将会展现更多更有挑战性的例子，也希望，这些例子足够刺激，能够满足我们的挑战欲。第一个例子是当我们说到混合策略时的一个经典例子：足球比赛中的点球。假设有两名参与者——守门员和踢点球的球员，现在我们允许有三种纯策略：踢球者会将球踢向左边、中间、右边，而同时守门员会扑向左边、右边或者守在中间。再以经典青少年片《黑色星期五》（*Friday the 13th*）为例，我们假设有三名参与者，一名杀手和两名青少年。少年会选择是否从前门或者后门逃出，而杀手会选择是埋伏在前门还是后门。最后一个例子印证了心理学家所提到的“旁观者效应”，此效应会在有固定参与者的情况下发挥作用。这不是一个可以用推理取胜的博弈，事实上，在纯策略中存在不对称的纳什均衡，只有在混合策略中才有对称均衡。

案例研究

足球比赛中的点球

经过120分钟的比赛，2006年世界杯决赛中法国和意大利进入了点球大战。一番点球激战后，在意大利4:3领先的情况下，意大利的发球者法比奥·格罗索（Fabio Grosso）面对在自家球门前严阵以待的法国门将法比安·巴特兹（Fabien Barthez）。格罗索起脚发球，球像火

箭般冲进了球门右角，而巴特兹扑向了左侧。意大利球员欢呼起来，而法国球员垂头丧气，脸上写了失望和难以置信。



2006 年世界杯决赛中点球大战

这种比赛，比的是守门员和发球者的心理。球员会向哪发球，而守门员又会往哪守球？这种情况不是足球比赛才有的，其他很多比赛都有相似的情形。在棒球比赛中，投手会投哪种球（快速直球、曲线球、擦边球等），而击球手又会往哪个方向击球。在橄榄球比赛中，防守者是否会突袭，进攻者是否会出其不意地进攻？在网球比赛中，球员是将球击到中线还是底线，而接球者应该选择哪个方向接球？这些都不仅仅是靠技术取胜的博弈比赛，还需要策略。

用足球比赛中的点球建模，假设发球者和守门员都无一例外的要决定到底选择哪边进攻或防守。从发球者的角度来看，选择的可能有：左、中、右（对于守门员来说，中间相当于原地不动）。对于发球者来说，其收益为得分，而守门员的收益相当于发球者不得分。收益矩阵如图 7—11 所示。

		守门员		
		左	中	右
发球者	左	0.65,0.35	0.95,0.05	0.95,0.05
	中	0.95,0.05	0,1	0.95,0.05
	右	0.95,0.05	0.95,0.05	0.65,0.35

图 7—11 点球大战

简单来说，胜利的频率和实际的操作有联系。如果守门员判断失误，比如当发球者向左边射球时，守门员扑向了右边，此时发球者得分的概率为 95%。但如果守门员判断正确了，发球者不得分的概率增加到 35%。双方同时选择中间是比较罕见的一种情形，而我们在这种情况下假设此时发球者肯定不能得分。

很明显，在这个博弈中没有纯策略纳什均衡。例如，如果巴特兹总是倾向选择右面，那么格拉索就会知道把球踢向中间或者左边，即如果巴特兹被预期扑向右面，格罗索就不会把球踢向右面。表 7—5 展示的是在发球者和守门员之间出现不同结果的具体频率，不出所料的是，他们都会倾向使自己的选择更具模糊性，因为这种不可预测性才是取胜的关键。

表 7—5 发球者与守门员的实际选择

发球者	守门员	发生概率（%）
左	左	19.6
左	中	0.9
左	右	21.9
中	左	3.6
中	中	0.3
中	右	3.6
右	左	21.7
右	中	0.5
右	右	27.6

发球者的混合策略是，选向左（用 k_l 表示），选向中（用 k_c 表示），选向右（用 k_r 表示）。可知 $k_l+k_c+k_r=1$，$k_c=1-k_l-k_r$，对于不同的 k_l, k_r 值，发球者有着混合策略。相似的，用 g_l, g_c, g_r 来表示守门员选择项左、中、右扑向的可能性，有 $g_c=1-g_l-g_r$，此时值 g_l，g_r 决定了守门员的混合策略。

一个纳什均衡包含了发球者随机化他的两种策略（左和右），或者是三者都随机化。我们设想下，每位参与者三种纯策略都为正概率是一个纳什均衡。那么这样来看，发球者向左、中、右发球的期望收益都是相同的。以下是各情况下不同的期望收益：

发球者向左的收益：

$g_l\times 0.65+（1-g_l-g_r）\times 0.95+g_r\times 0.95=0.95-0.3g_l$（7.5）

发球者向中的收益：

$g_l\times 0.95+（1-g_l-g_r）\times 0+g_r\times 0.95=0.95（g_l+g_r）$（7.6）

发球者向右的收益：

$g_l\times 0.95+（1-g_l-g_r）\times 0.95+g_r\times 0.65=0.95-0.3g_r$（7.7）

我们想为守门员找出一个策略——对 g_l 和 g_r 赋值，此时发球者所有的期望收益是相等的。算出等式（7.5）和（7.7），我们选择左和右的期望收益是一样的：

$0.95-0.3g_l=0.95-0.3g_r \Rightarrow g_l=g_r$

因此，在纳什均衡中，守门员向左和向右所产生结果的概率是一样的。换句话说，如果守门员均等的选择向右或者向左，那么守门员向左或向右去的概率是相同的。用 g 来表示这一相同的概率：$g_l=g=g_r$。在等式

(7.5)~(7.7)中，用 g 来替换 g_l，g_r，我们可以得出发球者选择要么向左或右的收益变为 $0.95-0.3g$，而选择中间的收益为 $1.9g$。而发球者想要使其三种选择的机会都均等，那么选择左（或右）和中的收益应该是相同的：

$$0.95-0.3g=1.9g \Rightarrow g=\frac{0.95}{2.2}\approx 0.43$$

总结起来，纳什均衡使守门员选择向左的概率为 0.43，选择向右的概率为 0.43，而选择留守中间的概率为 0.14。只有在满足这些概率的条件下，守门员才会均等地决定他的选择。

当然，守门员发现随机化他的三种选择才能使其选择最优，每种选择都需要相同的期望收益。同样的演绎方法可以推出发球者的策略以使门将对其选择的机会均等。这么一来就表明，发球者将球踢向左的概率为 0.43，踢向右的概率为 0.43，踢向中间的概率为 0.14。

有了这些均衡策略，球进门得分的概率可以计算如下：

$$0.43\times(0.43\times 0.65+0.14\times 0.95+0.43\times 0.95)+0.14\times(0.43\times 0.95+0.14\times 0+0.43\times 0.95)+0.43\times(0.43\times 0.95+0.14\times 0.95+0.43\times 0.65)=0.820\,44$$

第一项表示守门员扑向右面的概率为 0.43。在那种情况下，发球者也选择踢球到右面（发生的概率为 0.43），那么球进的概率为 65%，如果发球者选择中间（发生的概率为 0.14），那么进球的概率为 95%；如果发球者选择左边（发生的概率为 0.43），那么进球的概率为 0.95。第二项和第三项表示当守门员守中和扑向左。我们简单的理论可以预测进球的概率为 82%。

对此理论的另一个猜测是在众多选择中进球得分的概率是相等的。只有在得分机会相等的情况下，发球者才会乐意使选择踢球的方向随机化，即向左或者向右踢都行。一项研究给予相等化足够的支持：对于 22 个不同的发球者，按照发球者自然踢球的方向进球得分的概率为 82.68%，自然踢球方向的意思是一个习惯向左踢球的人会向左或中踢球，而一个习惯向右踢的人会向右或中踢球。而逆自然选择方向踢球（例如，习惯向右踢的人把球踢向左）而进球得分的概率为 81.11%。从数据上来说，82.68% 与 81.11% 两概率差异极小。所以按照预测的，无论把球踢向哪面，进球得分的概率是一样的。

鞭击：黑色星期五

在电影《黑色星期五》中，一位名叫杰森·沃里斯（Joson Voorhees）的男孩在 1957 年参加水晶湖野营时溺水而死。两名野营负责人在第二年被神秘地谋杀了。野营地在因此事关闭很多年后于 1979 年重新开放。一个乡村疯子警告说还会发生更多谋杀案，但没人相信。他能证实此事吗？杰森戴上一个曲棍球面具，悄悄拿上一把大而锋利的刀子，尾随野营的两位主管贝斯（Beth）和汤米（Tommy）到主楼。这两位主管能否生存下来或者杰森能否在主楼里得手？

贝斯和汤米分布于主楼里，他们能同时独自决定是从前门跑出亦或是从后门逃出。同样的，杰森也在决定着到底是在前门处埋伏还是躲在后门出口处。他们的最佳选择显而易见。对于贝斯和汤米最好的结果是他们两人从同一个门逃出，而杰森选择了另外一个门埋伏。这种结果的期望收益是当他们同时从前门跑出时为 3，同时从后门逃出时为 2，选择前门更好的原因是前门离主道路更近。对贝斯和汤米而言，另一较好的结果是三人同时选择同一门逃出或者潜伏。在二对一的情形下，逃跑仍然是可能的，此种情况下的收益为 0。

余下情况的结果由贝斯和汤米不同的选择可以计算得出。对汤米而言，次好的结果是贝斯和杰森选择同一出口，而他选择另外一个，汤米由此逃过杰森。最差的结果是汤米和杰森选择了同一出口（贝斯选择的是另一个），此种情况下，汤米的收益为-4。对贝斯而言，他的两种收益是刚好相反的。

最后我们分析杰森的情况，他最好的结果是他和贝斯（除开汤米）选择同一出口，或者他和汤米（除开贝斯）选择同一出口。杰森在这两种情形下的收益都是 2。杰森最差的结果是当汤米和贝斯都选择同一出口时，他选择的另外一个；此时他的收益为-2。对杰森而言，三人同时选择同一出口概率居中，此时收益为 0。

所有收益统计于图 7-12 中。汤米的表示于图横向上，贝斯的表示于纵向上，而杰森的用矩阵表示。三数字中第一个是汤米的，第二个是贝斯的，第三个是杰森的。这仅是一个部分能以智取胜的博弈，因为贝斯和汤米想胜过杰森，而杰森想胜过他俩，但贝斯和汤米都更倾向于选择同一出口。

在这里没有纯策略纳什均衡。为了推出在混合策略里的纳什均衡，用 t 表示汤米从前门跑出的概率，用 b 表示贝斯从前门跑出的概率，用 j 表示杰森选择前门的概率。先算杰森选择前门时的期望收益：

$$t\times b\times 0+t\times(1-b)\times 2+(1-t)\times b\times 2+(1-t)\times(1-b)\times(-2)=2+4t+4b-6tb \qquad (7.8)$$

杰森，前门

贝斯

		前门	后门
汤米	前门	0,0,0	-4,1,2
	后门	1,-4,2	2,2,-2

杰森，后门

贝斯

		前门	后门
汤米	前门	3,3,-2	1,-4,2
	后门	-4,1,2	0,0,0

图 7—12　黑色星期五

等式的含义如下：汤米和贝斯都试图从前门逃出的概率为 $t\times b$，如果杰森此时也在前门，他的期望收益应为 0。汤米选择前门而贝斯选择后门的概率为 $t\times(1-b)$，此时杰森可以攻击汤米，此时杰森的收益为 2。式子的第三部分表示的是汤米在后门而贝斯于前门的类似情况。最后 $(1-t)\times(1-b)$ 表示的是汤米和贝斯都从后门逃出，而杰森独自等在前门，此时他的收益为−2。由此杰森选择在后门埋伏时的支付如下所示：

$$t\times b\times(-2)+t\times(1-b)\times 2+(1-t)\times b\times 2+(1-t)\times(1-b)\times 0=2t+2b-6tb \qquad (7.9)$$

如果杰森随机化对前门和后门的选择，他的收益都是一样的。解式子（7.8）和（7.9）我们可以得到：

$$-2+4t+4b-6tb=2t+2b-6tb\Rightarrow t+b=1 \qquad (7.10)$$

因此，在纳什均衡下，如果杰森随机化选择而达到最优，那么汤姆选择前门和贝斯选择前门的概率之和应为 1。

回到汤米，他选择从前门跑出的期望收益为：

$$b\times j\times 0+(1-b)\times j\times(-4)+b\times(1-j)\times 3+(1-b)\times(1-j)\times 1=1+2b-5j=2bj$$

相反的，若他选择从后门逃出的期望收益为：

$$b\times j\times 1+(1-b)\times j\times 2+b\times(1-j)\times(-4)+(1-b)\times(1-j)\times 0=-4b+2j+3bj$$

对于汤米而言，若想使随机化达到最优，必须满足：

$$1+2b-5j+2bj=-4b+2j+3bj\Rightarrow 7j-6b+bj=1 \qquad (7.11)$$

因为汤米和贝斯的情形是对称的，所以贝斯的情形可以类似地推出如下：

$$7j-6t+tj=1 \qquad (7.12)$$

把式子（7.10），（7.11）和（7.12）联系起来，(t^*,b^*,j^*) 为纳什均衡的条件为：

$$t^*+b^*=1$$

$$7j^*-6b^*+b^*j^*=1$$

$$7j^*-6t^*+t^*j^*=1$$

如果上述条件成立，那么三位参与者都对其随机化选择是比较满意的。最后两个条件暗示 $t^*=b^*$。把此式与第一个联立，可得 $t^*+b^*=1$，由此可得 $t^*=\frac{1}{2}=b^*$。因此，汤米和贝斯疯狂冲出前门和后门的机会均等。因此，我们可以用（7.12）式推出杰森的均衡策略：

$$7j^*-6\times\frac{1}{2}+\frac{1}{2}\times j^*=1\Rightarrow j^*=\frac{8}{15}$$

杰森在前门的概率为 $\frac{3}{5}$。

那么汤米和贝斯同时逃走的概率是多少呢？也就是汤米和贝斯同时从前门逃出，杰森在后门的概率：

$$\left(\frac{1}{2}\right)\times\left(\frac{1}{2}\right)\times\left(\frac{7}{15}\right)=\frac{7}{60}$$

加上汤米和贝斯同时选择后门逃出，而杰森在前门的概率：

$$\left(\frac{1}{2}\right)\times\left(\frac{1}{2}\right)\times\left(\frac{8}{15}\right)=\frac{8}{60}$$

最后所得总和为 $\frac{15}{60}$，也即汤米和贝斯安然无恙逃出的概率为 25%——但即便他们都逃出去了，在黑色星期五里，总还有第二阶段的厄运等着他们。

旁观者效应

1964 年，纽约市有一段不为人知的经历，凯蒂·吉诺维斯（Kitty Genovese）在她公寓附近遇袭。尽管前后有 38 人听见了她的尖叫声，但没有一人上前帮助她。听着令人吃惊，但这种没有周围人回应的事也不是真的那么让人惊讶，因为这里存在旁观者效应。心理学家试验和理论研究表明当一个需要帮助的人处于集体环境而非个人环境时，旁人更不愿意伸出援助之手。所有听见凯蒂尖叫的人都知道还有其他人也听见了。研究还表明周围有更多人能提供帮助时，真正愿意提供帮助的人会更少。我们可以通过一些简单理论化的模型里的均衡策略行为看出以上现象的合理性。

假设有 $n\geq 2$ 的人面临着是否帮助那些需要帮助的人的决定。每人都自发地在置身事外与提供帮助间做选择。我们设想人们都有心相助并关心受害者是否得到帮助，并且帮助他们还有个人付出在里面。个人的收益有以下 4 种可能值，如图 7—13 所示。

		其他参与者的选择	
		全部置身事外	至少一人提供帮助
参与者的选择	提供帮助	a	c
	置身事外	d	b

图 7—13　旁观者博弈中参与者的收益

和以前我们讨论的博弈不同，我们不能准确地说明收益到底是多少，我们只能用字母 a,b,c,d 来表示数值意义，但必须满足 $a>d, b>c$。例如，$a=4, b=3, c=2, d=1$ 就符合条件，而 $a=5, b=3, c=4, d=1$ 就不行。这些对收益的判断到底暗示着什么偏好呢？$a>d$ 表示当其他人都不愿帮助受害人时，某人愿意帮助，而 $b>c$ 表示当有其他人愿意帮助时，某人不愿意再去帮助。人们其实愿意提供帮助，只是在有人已在帮助时不愿再麻烦了。

这个博弈中有 n 个纯策略纳什均衡，任一人选择提供帮助的同时，其他 $n-1$ 人都会选择置身事外。由于选择提供帮助的人是不同的，所以存在 n 个均衡。例如，考虑这样一种情形，有一个人选择提供帮助，然后 $2,\cdots,n$ 人选择置身事外。在其他 $n-1$ 人都置身事外的情形下，如果提供帮助，帮助者的收益为 a，如果不帮助，他的收益下降到 d，他如不帮助意味着受害者得不到任何帮助。所以帮助者 1 的策略是最优的。当周围其他的人也提供帮助时，他选择置身事外的收益为 b，而这会超过他再提供帮助的收益 c。这就是一个纳什均衡。当然，这些均衡在所有帮助的负担都落在一个人身上时是不对称的。在混合策略中推导出一均衡可以发现一个更为对称的解决方案。假设每人都有着 p 的概率去帮助。这一对称的均衡是在概率 p 的情况下实现，记为 p^*，即在有其他 $n-1$ 个人在 p^* 的概率下去帮助时，那么对其中任意一个人而言，在概率 p^* 时去帮助的策略最优。也就是说如果 $n-1$ 中的任意一个人在 p^* 概率下选择去帮助，那么其中另一个人在概率 p^* 下选择去帮助或不帮助都行，而且是均等的去选择。

倘若其他 $n-1$ 个人在 p 概率下选择帮助，一名帮助者选择帮助的期望收益如下：

$$\text{选择提供帮助的期望收益}=(1-p)^{n-1}\times a+[1-(1-p)^{n-1}]\times c \tag{7.13}$$

每个人在概率 $1-p$ 下选择置身事外，因为他们的选择是独立的，所以 $n-1$ 人都置身事外的概率为 $p-1$ 的 $n-1$ 次方，即 $(1-p)^{n-1}$。在没有其他人选择提供帮助的情况下，一个人选择提供帮助的收益为 a。在概率 $1-(1-p)^{n-1}$，当其他 $n-1$ 个人中至少有一人提供帮助时，他所得的收益为 c。[①]这就解释了（7.13）所示的期望收益。

那么此人选择置身事外的期望收益又是多少呢？

$$\text{选择置身事外的期望收益}=(1-p)^{n-1}\times d+[1-(1-p)^{n-1}]\times b \tag{7.14}$$

同样，倘若在混合策略下，其他所有人置身事外的概率是 $(1-p)^{n-1}$。如果这个人同样选择置身事外，那么他的收益就是 d，因为他会因受害人被忽视而感到难过。在概率 $1-(1-p)^{n-1}$ 下，只要在其他任何一人选择提供帮助的情况下，他置身事外的收益就是 b。

这个均衡策略 p^* 是由值 p 解式（7.13）（7.14）后得出的，所以这个人愿意随机化其选择：

$$(1-p^*)^{n-1}a+[1-(1-p^*)^{n-1}]c=(1-p^*)^{n-1}d+[1-(1-p^*)^{n-1}]b$$

进行一系列代数运算后，我们可以解出对称的均衡混合策略：

$$(1-p^*)^{n-1}a+[1-(1-p^*)^{n-1}]c=(1-p^*)^{n-1}d+[1-(1-p^*)^{n-1}]b$$

$$(1-p^*)^{n-1}a+c-(1-p^*)^{n-1}c=(1-p^*)^{n-1}d+b-(1-p^*)^{n-1}b$$

$$(1-p^*)^{n-1}[(a-d)+(b-c)]=b-c$$

$$(1-p^*)^{n-1}=\frac{b-c}{(a-d)+(b-c)}$$

$$[(1-p^*)^{n-1}]^{\frac{1}{n-1}}=[\frac{b-c}{(a-d)+(b-c)}]^{\frac{1}{n-1}}$$

$$1-p^*=[\frac{b-c}{(a-d)+(b-c)}]^{\frac{1}{n-1}}$$

$$p^*=1-[\frac{b-c}{(a-d)+(b-c)}]^{\frac{1}{n-1}} \tag{7.15}$$

为了分析这个均衡等式的特性，假设 $a=4, b=3, c=2, d=1$，那么等式表达就是：

$$p^*=1-(\frac{1}{4})^{\frac{1}{n-1}}$$

均衡等式概率曲线在横轴标出的人数上描绘出，见图 7—14。当有更多的人选择去帮助的时候，其中一人去帮助的概率就会降低，这也就是先前心理学家分析的心理特性。

①请注意受害者要么得到 $n-1$ 人中至少1个人的帮助，要么得不到任何帮助。因此，发生这两种事件的概率之和为1。后者的发生的概率为 $(1-p)^{n-1}$，则前者的概率为 $1-(1-p)^{n-1}$。——作者注

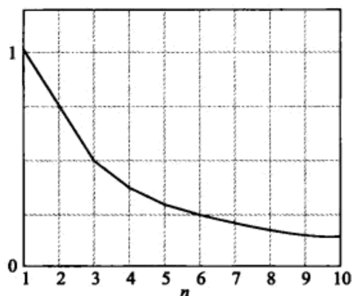


图 7—14　有一个人提供帮助的概率

尽管每个人都更不愿提供帮助，但有更多的人可以提供帮助。这样一来，在一个有更多人的群体里，是否会有至少一个人愿意提供帮助是很不确定的。为了解决这个问题，我们只需要计算出 n 个人中的一个在概率 $1-(1-p^*)^n$，或概率 $1-\left(\frac{1}{4}\right)^{\frac{n}{n-1}}$下愿意提供帮助，如图 7—15 所示。让人惊奇的是，更多的人使受害人得到帮助的可能性更低。

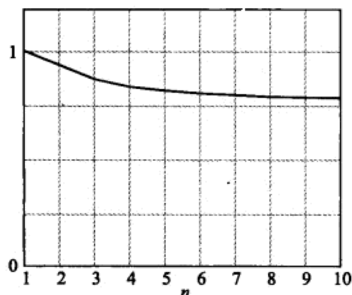


图 7—15　至少有一个人提供帮助的概率

7.5 纯冲突和防范行为博弈

回顾先前的二人常和博弈是关于二人收益总和总为常量的博弈。因为常和这一属性暗示了一参与者收益的提高总会降低另一人的收益，常和博弈是纯冲突的情况。换句话说，对一人有好处时必定使另一人受损。

考虑福尔摩斯和莫瑞特教授在阿瑟的作品《最后的决策》(*The Final Solution*) 里所面临的境遇，福尔摩斯给莫瑞特的非法活动造成了非常严重，或许是不可弥补的损失，而莫瑞特则急切的寻求复仇。福尔摩斯最初的逃跑路线是搭乘火车由伦敦到多佛尔，然后从那到欧洲大陆。当火车启动时，他发现莫瑞特就在火车站台。福尔摩斯及时向他的顾问寻求建议，顾问的建议是让他搭乘一辆更安全的火车离开。福尔摩斯面临两种选择，要么选择搭车去多佛尔，要么在路途中间的坎特伯雷就下车。莫瑞特很聪明，他了解所有选择的可能性，他也面对同样的选择。

福尔摩斯认为如果和莫瑞特在站台相遇，他很有可能被莫瑞特杀害。如果他能平安无事地到达多佛尔，他就能顺利逃离莫瑞特的视线。即便莫瑞特猜中福尔摩斯的想法，他仍坚持选择去多佛尔，因为他能更好地逃离去大陆。这个博弈的策略图在图 7—16 中可见。需要注意的是这是个纯冲突策略，收益总和是 100。

		莫瑞特	
		多佛尔	坎特伯雷
福尔摩斯	多佛尔	20,80	90,10
	坎特伯雷	70,30	10,90

图 7—16　最终解决方案

如果你已经阅读了许多关于福尔摩斯的故事，你就会发现他是个聪明且自大的人。在处理这个策略性情形时，采取一个保守的方法并不符合他的性格，因为他足够聪明且知道莫瑞特会是一个很强的对手。所以与其猜测福尔摩斯会先算到莫瑞特的选择后再出对策，不如让我们思考下福尔摩斯会在选择策略上采取一个更保守的路线。

我们假设福尔摩斯相信不管他采取什么策略，莫瑞特总会预测出来，并减少福尔摩斯的期望收益。这样，福尔摩斯预测自己减少的收益反映出他在这种情形下的悲观，但值得注意的是，既然这是个常和博弈，莫瑞特减少福尔摩斯收益相当于增加自己的收益。已经知道会对莫瑞特所做的反应后，福尔摩斯会选择混合策略以增加自己的期望收益。换句话说，他通过采取防范措施来使自己的悲观情绪得到改善。我们刚对福尔摩斯所选择的策略的描述即我们所说的**极大极小策略**（maxmin strategy）。更通俗地说，极大极小策略即用来使参与者达到最优状态的策略，即会设想到其他参与者会通过选择策略来使参与者的收益最小。

为了推出福尔摩斯的极大极小策略，需要运用一些记号。用 $V_H(p_H, s_M)$ 表示福尔摩斯的期望收益，如果他选择混合策略 p_H（p_H 即表示他选择去多佛尔的概率），则用 s_M 来表示莫瑞特的纯策略（即表示莫瑞特或去多佛尔或去坎特伯雷）。福尔摩斯的极大极小策略可以解

决以下问题：在已知 p_H 情况下，当 s_M 被选择来使 V_H（p_H,s_M）最小化时，用 p_H 来使 V_H（p_H,s_M）最大化。如果我们用 b（p_H）来表示 s_M 使 V_H（p_H,s_M）最小化的值，那么福尔摩斯的问题可以表述为怎么样定 p_H 的值来使 V_H（p_H,s_M）最大化。因此，在福尔摩斯选择 p_H 时，他会料想到莫瑞特会用 b（p_H）来使他的期望收益最小化。不管福尔摩斯怎么做，莫瑞特总会预料到并作出相应的决策。

为了解决这个问题，首先考虑当莫瑞特选择去多佛尔时，福尔摩斯的期望收益：

V_H（p_H, 多佛尔）$=p_H\times 20+(1-p_H)\times 70=70-50p_H$

如果莫瑞特选择去坎特伯雷，福尔摩斯的期望收益：

V_H（p_H, 坎特伯雷）$=p_H\times 90+(1-p_H)\times 10=10+80p_H$

所有可能出现的期望收益 p_H 已经在图 7—17 中标出。回顾下我们的猜想，不管福尔摩斯选择的 p_H 值是多少，莫瑞特总会算出来并选择相应的车站下车来使他的期望收益最小。结果,福尔摩斯面临所示两条线的较低的一条，图中加粗的线。例如，如果福尔摩斯把 p_H 值设为 $p_H=\frac{1}{4}$，那么因为莫瑞特选择坎特伯雷会使他的收益最小化，所以他所预料的期望收益就应为 $10+80\times\frac{1}{4}=30$。

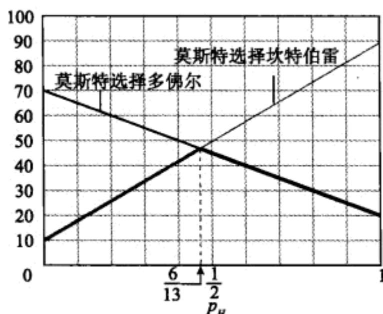


图 7—17　福尔摩斯的极大极小解决方法

在图 7—17 中存在一个对于福尔摩斯来说的极大点，即两条线的交点。我们可以通过如下的计算来得出他的极大极小策略：

$$70-50p_H=10+80p_H \Rightarrow p_H=\frac{6}{13}$$

在假设福尔摩斯无论采取何种混合策略时，莫瑞特总会采取相应的措施来使福尔摩斯的期望收益最小化，因此，$p_H=\frac{6}{13}$是能使福尔摩斯期望收益最大化的混合策略。$p_H=\frac{6}{13}$之所以叫做福尔摩斯的极大极小策略是因为它能使他的最小收益最大化，正如图 7—17 所描绘的那样。

再继续分析之前，我们注意到防范意识很强的福尔摩斯严格的倾向随机化选择，这是非防范纳什均衡里没有的案例。如果一参与者在纳什均衡里选择随机化——例如，关于多佛尔和坎特伯雷两种选择，那么他来自选择混合策略的期望收益与他选择单一的纯策略如多佛尔或坎特伯雷，或其中任意一个纯策略的期望收益是一样的。在纳什均衡下，一参与者选择一策略，而另一位参与者的策略就由此而固定了。这就和可以找出极大极小解决方法的防范选择的情形不太一样了。福尔摩斯不认为莫瑞特的选择是固定的；相反，他更期望莫瑞特能选择一混合策略以使他的收益最小化。如果一名参与者知道你会做出的策略，那么你能作出随机化选择将会更好。

如果这点上我们还不能完全清楚，那假设福尔摩斯选择去多佛尔的情景。莫瑞特选择去多佛尔，福尔摩斯的期望收益为 20。如果福尔摩斯选择去坎特伯雷，那么为了达到最小化他收益的目的，莫瑞特会选择去坎特伯雷，此时福尔摩斯的收益为 10。现在考虑福尔摩斯轻掷一枚硬币，那么去多佛尔或者坎特伯雷的概率都是 50%。从图 7—17 中可以看出，莫瑞特选择去多佛尔（当 $p_H=\frac{1}{2}$）时，福尔摩斯的收益为 $70-50\times\frac{1}{2}=45$。因此福尔摩斯随机化选择的结果要优于单纯的纯策略的选择。随机化使得莫瑞特到多佛尔抓住福尔摩斯的概率为 50%，此时福尔摩斯的收益为 20。但同时福尔摩斯有 50% 的概率去坎特伯雷，这样一来他逃过莫瑞特的收益达到了 70。一些不确定的预测在这里是很必要的，而混合策略在这里开始起作用。

已经解决了福尔摩斯的极大极小策略，现在思考倘若莫瑞特预测福尔摩斯要么选择去多佛尔要么去坎特伯雷以达到最小化他收益的举动，他同样具有更强的防范意识，所以他也会选择混合策略，用 p_M（他选择去多佛尔的概率）来表示。那么如果福尔摩斯去多佛尔，莫瑞特来自选择策略 p_M 的期望收益为：

$$p_M\times 80+(1-p_M)\times 10=10+70p_M$$

以及如果福尔摩斯去坎特伯雷，其值为：

$$p_M\times 30+(1-p_M)\times 90=90-60p_M$$

把上面的式子在图 7—18 中描绘出来，莫瑞特的期

望收益（在福尔摩斯会随机选择多佛尔或坎特伯雷以使莫瑞特收益尽可能小的条件下）为图中加粗的线，则可知极大极小策略值 $p_M=\frac{8}{13}$。

总结得出，如果福尔摩斯和莫瑞特并不如他们看起来的那么骄傲自大，而是如我们假想般的心思缜密且具有很强防范意识的话，那么福尔摩斯选择去多佛尔的概率为$\frac{6}{13}$，而莫瑞特去多佛尔的概率为$\frac{8}{13}$。这就是这个博弈的极大极小解决方法。

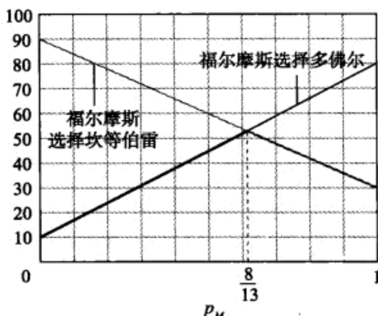


图 7—18　莫斯特的极大极小解决方法

现在让我们思考一个有趣的部分。假设我们不是去找出混合策略里的纳什均衡，即在知道福尔摩斯能准确判断莫瑞特的混合策略前提下，他会选择混合策略来使自己的收益最大化。同样，莫瑞特在能知道福尔摩斯的混合策略的条件下，他也会选择混合策略以使自己收益最大化。那么，用第 6 章所学的方法，我们就可以看出这个独特的混合策略纳什均衡使得福尔摩斯去多佛尔的概率为$\frac{6}{13}$，使莫瑞特去多佛尔的概率为$\frac{8}{13}$。例如，在已知莫瑞特用此混合策略时，福尔摩斯选择去多佛尔的期望收益为：

$$\left(\frac{8}{13}\right)\times 20+\left(\frac{5}{13}\right)\times 90=\frac{610}{13}$$

选择去坎特伯雷的收益为：

$$\left(\frac{8}{13}\right)\times 70+\left(\frac{5}{13}\right)\times 10=\frac{610}{13}$$

由此可见福尔摩斯对此两种纯策略的选择以及这之间的任意随机化选择是无所谓的。随之得出的结论是，针对莫瑞特在$\frac{8}{13}$的概率下去多佛尔的情况下，福尔摩斯在$\frac{6}{13}$的概率下去多佛尔的选择是最佳的，同样可以得出莫瑞特的最佳选择亦是如此。

那么，注意到了纳什均衡和极大极小解决方法是一样的。参与双方表现出防范举动会在能正确预测另一方会怎么做的前提下产生最佳的结果。换句话说，防范不需要代价，因为当参与一方这样做时，是对另一方做法的最优反应。这并不是巧合。



点拨

极大极小的性质：对于任何纯冲突的两个参与者而言，极大极小解决方法即是纳什均衡。更进一步说，如果纯冲突博弈双方有着随机化选择的纳什均衡（如他们不用纯策略），那么每一方的纳什均衡都是一个极大极小策略。

上面关于极大极小性质的说法归功于 20 世纪一名伟大的数学家约翰・冯・诺伊曼。他的贡献远不止纯数学方面，更包括了物理和计算机科学。特别让我们由衷感谢的是，他在 1944 年和奥斯卡・摩根斯顿共同撰写了名为《博弈论与经济行为》的博弈论经典著作。

本章总结与练习题

本章总结

这章之前，我们为博弈中的一名参与者建模，决定他在纯策略的选择中应该如何抉择。在一些背景中，对于随机参与概率的考虑是很自然的。为了获得这一选择，随机版本的博弈由此发展而来。对于每位参与者的策略来说这属于混合策略，这给每个参与者的纯策略赋予一概率。为了总体评估混合策略，参与者需要从混合策略暗示的概率和原博弈中的收益来计算出他的期望收益。随机版本博弈吸引我们的一点是他总能有一个解决途径；也就是说在这样的混合策略中总存在纳什均衡。

我们研究了在随机化选择尤为突出的两个经典博弈。第一个是参与者想要按对方没有预料的方式做出行动。因为参与者总是想使自己变得不可预测，随机化的选择一纯策略是解决此问题的途径之一。以智取胜的博弈例子有许多，如体育方面的（点球），犯罪方面的（警车巡逻和毒贩交易），战争方面的（阿维让切斯防线），夏日野营方面的（黑色星期五）。

第二个是挤出效应的对称博弈论——即当有更多人参与选择某一策略时，该策略的吸引力更小。这章中这方面的例子是市场的进入博弈和旁观者博弈。尽管挤出效应的博弈会有不对称的纯策略纳什均衡，但他们基本上不会有对称的纯策略纳什均衡。对于对称博弈而言，对称的解决方法是有用的，而且这些博弈通常有混合策略里的对称纳什均衡在里面。

在纳什均衡里，一个参与者在随机化选择中使用混合策略是很有必要的，其对纯策略选择的可能性相同，所有被赋予正的概率的纯策略肯定会产生最高的期望收益，这也就是为什么参与者愿意让随机机制来决定他的选择行为。如果不那样的话，当在随机化实施之后，这名参与者很可能忽视潜在的选择而去选择一个更好的。机会均等的均衡性质在解决混合策略纳什均衡中是极为有用的。

参与者的极大极小策略可以定义为当他做出极高防范措施时，能料想到无论他作出何种选择，对方总会作出相应的行为以使自己的收益最小化时，此时他自己的策略可以使他自己的收益最大化。在两人参与的纯冲突博弈中，在两人都用极大极小策略时，即为纳什均衡。因此，无论此参与者是乐观（认为他能预料到对方的决策）还是悲观（认为对方会比他更聪明），合理的选择行为都一样。

练习题

1. 图 7—19 中的电话游戏来自 4.2 小节。在此混合策略中找出所有的纳什均衡。

		温妮	
		打电话	等电话
科琳	打电话	0,0	2,3
	等电话	3,2	1,1

图 7—19　打电话博弈

2. 还有 3 球和 2 次进攻，所有垒是空的。击球手想要使击中或跑垒的概率最大化，而投手想使其概率最小。投手在选择掷一快速球还是曲线球之间做决定，而击球手在准备接快速球和曲线球之间做决定。每人的策略选择如图 7—20 所示。找出此混合策略中的所有纳什均衡。

		投手	
		快速球	曲线球
击球手	快速球	0.35,0.65	0.3,0.7
	曲线球	0.2,0.8	0.5,0.5

图 7—20　棒球

3. 现在是春假，你在 95 号洲际公路上向南行驶，目的地为福特劳特达尔。你是在限制的 65 英里每小时以内行驶呢，还是你希望加速到 80 英里每小时？那州际警察又会如何呢？他们是设置限速标志？还是不管这个直接开进市区在克里斯皮克莱米寻找是否有写着“鲜而热”字样的发光霓红灯招牌呢？警察通常喜欢去抓那些超速行驶的人，但同时他们又不会设置限速牌，不然就没人可抓了。此策略的表格如图 7—21 所示。驾驶员要么合法的按 65 英里每小时行驶，要么加速到 80 英里每小时。警察要么设置限速牌，要么冲进市区抓住那些超速行驶的驾驶员。驾驶员最好的结果是超速但没被抓住，此时她的收益为 70；最差的结果是她超速，且被抓了，此时她的收益是 10；如果她合法行驶，则收益为 40，此时警察是否参与没有影响。警察最好的结果是他设置了限速牌并抓住了超速者，此时他的收益为 100；而其最差的结果是他设置了限速牌而且没有开出一张罚单，此时他的收益为 20；当他选择去克里斯皮克莱米时，其收益为 50。找出此混合策略中的所有纳什均衡。

		州际警察	
		设置限速牌	克里斯皮克莱米
司机	80英里每小时	10,100	70,50
	65英里每小时	40,20	40,50

图 7—21　限速牌与超速

4. 一路贼和一受害者在一条漆黑的街道相遇。这

个路贼早就在思考是否应该带把枪，如果带了，是否在抢劫过程中使用。如果路贼没有使用枪（要么因为他没有枪，要么因为他把枪藏了起来），然后受害者再来决定到底要不要反抗。（注意如果路贼有枪并且亮出枪来，那么因为受害者的策略是在没有枪的情况下怎么做，所以无论受害者作出何种决策他的收益都是 5）。策略形势表如图 7—22 所示。注意所有的收益都是特定的，除了路贼选择持枪并使用时路贼的收益。找出在满足 x 值条件下时路贼随机化其两种纯策略（有枪但不拿出使用和无枪）以及受害者在随机化其两种选择（反抗和不反抗）的纳什均衡。

		受害人	
		反抗	不反抗
路贼	无枪	2,6	6,3
	有枪但不拿出使用	3,2	5,4
	持枪并使用	x,5	x,5

图 7—22

5. 在 7.4 小节的点球决赛中，找出在发球者仅随机化两种纯策略，即左和右，因此导致中间的概率为 0 时的纳什均衡。

6. 在图 7—23 所列出的博弈中，找出所有的混合策略纳什均衡。

		参与者1		
		x	y	z
参与者2	a	2,3	1,4	3,2
	b	5,1	2,3	1,2
	c	3,7	4,6	5,4
	d	4,2	1,3	6,1

图 7—23

7. 如图 7—24 所示，在混合策略中找出所有的纳什均衡。

		参与者2		
		左	中	右
参与者1	上	2,2	0,0	1,3
	中	1,3	3,0	1,0
	下	3,1	2,3	2,2

图 7—24

8. 在橄榄球的比赛的最后阶段，橄榄球的球队落后一分，球队决定再进两分，如果成功，比赛胜利。他们进攻时面临三种选择：往左，往右和往中间发起进攻。防御方在考虑应该是防守两翼还是中间。防御方的收益是进攻方不得分，而进攻方的收益是得分。找出所有的混合策略纳什均衡。（见图 7—25）

		进攻		
		向左进攻	向中间进攻	向右进攻
防守	防守两翼	0.6,0.4	4,0.6	0.6,0.4
	防守中间	3,0.7	0.5,0.5	3,0.7

图 7—25 橄榄球比赛的两分

9. 图 7—26 所示的是小孩常玩的游戏剪刀、石头、布。证明每位参与者在赋予其纯策略相等概率时是一对称的纳什均衡。

		丽萨		
		石头	布	剪刀
巴特	石头	0,0	−1,1	1,−1
	布	1,−1	0,0	−1,1
	剪刀	−1,1	1,−1	0,0

图 7—26 剪刀、石头、布

10. 三位参与者的每一位都要在纯策略走还是停之间做决定。选择走的收益为 $120/m$，m 表示选择走的人数，选择停的收益为 55（这与其他参与者怎么选择无关）。在此混合策略中找出所有的纳什均衡。

11. 总数大于等于 2 的公司在考虑是否进入市场，进入的成本是 30。如果只有一家公司进入，毛利润是 200。如果有超过一家的公司选择进入，则每位进入者的毛利润是 40，公司进入的收益是进入后的毛利润减去进入成本，而选择不进入的收益为 60。在此混合策略中找出一个对称的纳什均衡。

12. 萨达姆•侯赛因在决定当联合国在决定去寻找阿富汗的大规模杀伤性武器时，他应该把这些武器放哪。他成功藏匿这些武器的收益为 5，被联合国发现这些武器的收益为 2。对联合国来说，找到这些武器的收益为 9，未能发现的收益为 4。侯赛因可以把这些武器藏于建筑 X，Y 或 Z 中。联合国将决定去检查哪个建筑物，因为受时间和人力的限制，他们不能检查所有建筑物。

a. 假设联合国有两种纯策略：他们要么检查建筑 X 和 Y（这两建筑物的地理位置相隔较近），要么检查建筑物 Z。在此混合策略中找出一个纳什均衡。

b. 假设联合国可以检查任意两座建筑物，则其面临 3 种纯策略。联合国可以查 X 和 Y，X 和 Z，Y 和 Z。在此混合策略中找出所有的纳什均衡。

附　录

混合策略中纳什均衡的正式定义

考虑任何有限博弈——即有有限人数的参与者和每人所持策略有限的博弈。如果参与者 i 有 m_i 个纯策略，那么将她所有纯策略表示为$\{s_i^1,\cdots,s_i^{m_i}\}$。现在在这个博弈中考虑它的随机版本以使一混合策略赋予每位参与者的纯策略相等的概率。用p_i^j表示参与者 i 的纯策略s_i^j的概率。混合策略有 m_i 个数值，每个值介于区间 [0,1] 之间，总和为 1。换句话说，只要满足下式，$(p_i^1,\cdots,p_i^{m_i})$ 就为一混合策略：

$$p_i^j \geqslant 0 \qquad j=1,2,\cdots,m_i \text{ 且 } p_i^1+\cdots+p_i^{m_i}=1$$

参与者 i 的混合策略包含所有在混合策略 $(p_i^1,\cdots,p_i^{m_i})$ 中的所有值并满足进一步的条件。用 p_i 表示参与者 i 的混合策略。记住纯策略是混合策略的一种特殊情况，所以我们已经加上有随机化可能的值。

另一种关键的博弈的随机版本是一参与者用他的期望收益来评估他的混合策略。让我们用通常的两人参与的博弈来定义。当参与者 1 选择纯策略 s_1^j，参与者 2 选择纯策略 s_2^k 时，用 $V_1(s_1^j,s_2^k)$ 来表示他的收益。有了这个表达式，当参与者 1 用混合策略 $(p_1^1,\cdots,p_1^{m_1})$，参与者 2 选择混合策略 $(p_2^1,\cdots,p_2^{m_2})$ 时，我们可以定义她的期望收益为：

$$\sum_{j=1}^{m_1}\sum_{k=1}^{m_2} p_1^j \times p_2^k \times V_1(s_1^j,s_2^k)$$

参与者 2 的期望收益为：

$$\sum_{j=1}^{m_1}\sum_{k=1}^{m_2} p_1^j \times p_2^k \times V_2(s_1^j,s_2^k)$$

当参与者在做使用混合策略时随机化，那么两个随机事件（由参与者 1 选择的纯策略和参与者 2 选择的纯策略）相互独立。那就是为什么参与者 1 使用策略 s_1^j 和参与者 2 使用策略 s_2^k 的概率为 $p_1^j \times p_2^k$。

有了以上概率，我们就能正式定义混合策略中纳什均衡的条件了。在两人博弈中，如果满足以下两个条件，混合策略组 $(\hat{p}_1^1,\cdots,\hat{p}_1^{m_1})$ 和 $(\hat{p}_2^1,\cdots,\hat{p}_2^{m_2})$ 就是一纳什均衡：

$$\sum_{j=1}^{m_1}\sum_{k=1}^{m_2} \hat{p}_1^j\ \hat{p}_2^k V_1(s_1^j,s_2^k) \geqslant \sum_{j=1}^{m_1}\sum_{k=1}^{m_2} p_1^j\ \hat{p}_2^k V_1(s_1^j,s_2^k) \qquad (p_1^1,\cdots,p_1^{m_1}) \in P_1 \tag{7.16}$$

$$\sum_{j=1}^{m_1}\sum_{k=1}^{m_2} \hat{p}_1^j\ \hat{p}_2^k V_2(s_1^j,s_2^k) \geqslant \sum_{j=1}^{m_1}\sum_{k=1}^{m_2} \hat{p}_1^j\ p_2^k V_2(s_1^j,s_2^k) \qquad (p_2^1,\cdots,p_2^{m_2}) \in P_2 \tag{7.17}$$

等式（7.16）是确保参与者 1 混合策略最优化的条件，等式（7.17）是确保参与者 2 混合策略最优化的条件。例如，（7.16）式表明，若参与者 2 选择策略 $(\hat{p}_2^1,\cdots,\hat{p}_2^{m_2})$，没有其他更好的策略能使她的期望收益大于参与者 1 选择策略 $(\hat{p}_1^1,\cdots,\hat{p}_1^{m_1})$ 时的期望收益。

第三部分　求解扩展式博弈

第8章

交替：完美信息下的序列博弈

学习目标

8.1 回顾归纳法及子博弈精炼纳什均衡
8.2 等待博弈：抢先博弈与消耗战
8.3 人类在理智地使用逆向归纳法吗

本章导读

前几章的重点旨在探讨策略式博弈，现在我们将开始讨论扩展式博弈，图8—1是对第2章绑架博弈的再现。这里我们对诸多扩展式博弈进行回顾，要记住，图8—1的每一个黑点代表一个决策节点，在每一个节点处，参与者都要从此节点下属的几个分支选项中做出选择。在完全信息博弈，也就是本章我们将要讨论的博弈中，对每一个参与者而言，图8—1中每一个决策节点就意味着一个策略，一种行为。因此，对盖伊而言，他有三种行为可选，因为他有三个决策节点，同理，薇薇卡只有一个策略，包含一种行为。

纳什均衡的概念可以直接应用到这种扩展式博弈中，然而，如果我们首先得到他的策略情境，然后再探寻适合他的纳什均衡将会更加简单。因为用于扩展式博弈中的纳什均衡与用于策略形式中的纳什均衡是一致的，所以这种方法也是合乎情理的。

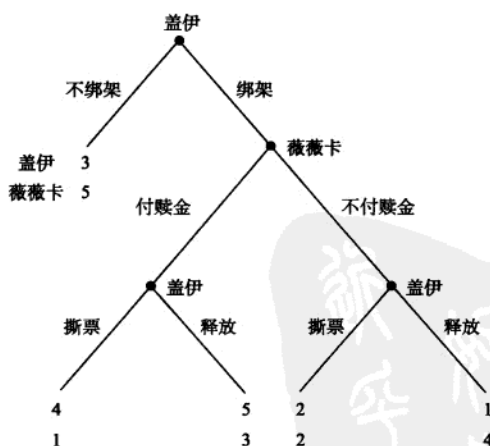


图8—1 扩展式博弈中的绑架博弈（源于第2章）

为了继续研究出一种方法，在图 8—2 中我们用一种与图 8—1 非常接近的策略情境进行表述。下面我们描述一下盖伊的策略，我们认为他的第一个行动（绑架或者不绑架）源于最初的节点，第二个行动是在薇薇卡交完赎金后决定的（撕票或者释放），第三步行动是在薇薇卡拒交赎金的情况下决定的（撕票或者释放）。

薇薇卡（受害人的妻子）

盖伊（绑架者）		付赎金	不付赎金
	不绑架/撕票/撕票	3,5	3,5
	不绑架/撕票/释放	3,5	3,5
	不绑架/释放/撕票	3,5	3,5
	不绑架/释放/释放	3,5	3,5
	绑架/撕票/撕票	4,1	2,2
	绑架/撕票/释放	4,1	1,4
	绑架/释放/撕票	5,3	2,2
	绑架/释放/释放	5,3	1,4

图 8—2　绑架博弈的策略形式

逐一研究这 16 种策略组合，我们发现存在 5 个纯策略形式的纳什均衡。（绑架 / 释放 / 撕票，付赎金）是其中之一，假设奥兰多被绑架，薇薇卡支付赎金，接着奥兰多被释放。这样，在每一个不发生绑架的假设中就存在 4 种纳什均衡：（1）（不绑架 / 撕票 / 撕票，不支付赎金）；（2）（不绑架 / 撕票 / 释放，不支付赎金）；（3）（不绑架 / 释放 / 撕票，不支付赎金）；（4）（不绑架 / 释放 / 释放，不支付赎金）。每一个这样的案例假设中，盖伊都不会绑架奥兰多，因为他的家属没有支付赎金的意图。上述的这 4 种均衡只有在下面这种情况中会有所不同：随后的博弈中，盖伊的策略让他做出绑架的选择，而这种行为如果源于纳什均衡是不会发生的。

对于绑架博弈而言，纳什均衡不是很恰当。这一策略预测绑架行为是否会发生；也就是说，一些均衡让奥兰多遭绑架，而其他的没有。但是所有的均衡都是一样的吗？实际上，这 4 种均衡中存在的一些难题会避免绑架行为的发生。薇薇卡做出了一个不太可信的恐吓，而盖伊却信以为真。盖伊不会绑架奥兰多，因为他确信薇薇卡不会付赎金。然而，只有在盖伊不会进行绑架时，薇薇卡的策略才是最优的，所以不管她的策略如何，收益是相同的。但是如果假设与盖伊的策略相反，他确实绑架了奥兰多，薇薇卡会坚持不付赎金吗？她很清楚盖伊的偏好 —— 如果不付赎金，盖伊倾向于撕票（收益为 2，释放奥兰多的收益为 1），而如果盖伊收到了赎金，他更倾向于释放奥兰多（收益为 5，撕票的收益为 4）。因此，似乎付赎金释放奥兰多能创造最大收益。现在我们认为盖伊知道所有信息，因此不管薇薇卡说什么他都会按原计划行事，他应该相信如果奥兰多被绑架薇薇卡将付赎金，因为这是令盖伊释放奥兰多的唯一方法。

综上所述，这 4 个维持"不绑架"状态的纳什均衡不是很有说服力。基于盖伊的想法，他们被断定一旦绑架事件发生，薇薇卡将违反常规不付赎金。但是薇薇卡只是虚张声势，而盖伊会信以为真。引发绑架事件的唯一的均衡不会让薇薇卡做出违反常规的行为，也不会令盖伊相信她将做出违反常规的行为。这样我们更容易得出这样的论断，最令人信服的纳什均衡是（绑架 / 释放 / 撕票，付赎金）。

本章及以后的章节将集中探讨扩展式博弈的问题，特别是我们将得出一些纳什均衡，这些均衡通过相信另一个参与者会做出违反常规的行为，或者某种被称为不可置信的威胁的博弈理论（incredible threat, 在这里，"不可置信"的意思是"没办法相信"，而不是"令人敬畏"）而得以维系。本章将重点探讨完全信息博弈，一个可以帮助我们脱离绑架博弈中这 4 种纳什均衡的解决方案。我们还将学习众所周知的子博弈精炼纳什均衡。

8.1 逆向归纳法及子博弈精炼纳什均衡

回顾前面的内容，我们知道一个策略就是一个非常规的规则，对于每个参与者而言，这个规则指出了源于每一个信息的行动。纳什均衡要求其他参与者的策略为既定策略时，每个参与者的策略都是最优的。当然，在每一个决策节点，策略所需的条件能否得到满足不会影响参与者的收益；当某个参与者被要求执行这一行动时，偶然事件下的行动就会被提出。因此，纳什均衡不要求所有意外情况下既定的行为是最优的，只要求这些行为满足均衡博弈的要求（当参与者使用他们的均衡策略时，就会发生一连串的博弈）。

为了让这一点更加明确，我们需要考虑这个策略的大体情况（绑架 / 释放 / 撕票，付赎金）。均衡博弈包含绑架事件的发生，在这一案例中，薇薇卡的策略会为她指出应对绑架事件的最佳行动（付赎金）。尽管这是一个偶然事件，但薇薇卡付赎金的情形还是会出现，此时，盖伊必定会制定出最佳策略（释放奥兰多）。现在我们用另外一个策略（不绑架 / 撕票 / 撕票，不付赎金）与这一策略进行对比，这一策略中，均衡博弈预测到不会有绑架事件的发生。薇薇卡的策略指出了绑架事件发生后其应该采取的行为，这一策略不需要研究出应对此事件的最佳行为，因为在最初的节点处，当盖伊的策略没有绑架行为时，绑架事件是不会发生的。我们注意到薇薇卡的收益是不受其策略影响的（如果你存在怀疑可以进行验证。）类似的，在绑架事件中，盖伊的策略不需要一个最佳决策而薇薇卡付了赎金，尽管这种偶然几乎不会发生。

关键在于纳什均衡并不需要一个为这些决策节点制定出最佳行动的策略，而均衡博弈跟这些节点无关。如果一个特别的偶然事件从来都未被意识到，那么应对这一偶然事件的策略对参与者而言是无足轻重的。纳什均衡真正需要的是在均衡博弈中所有决策节点处，让每个参与者的策略都能制定出最佳行动。

纳什均衡只有在受到某种因素制约的情况下才是最优的。**子博弈精炼纳什均衡的思想旨在将既定的参与者的最佳行动扩展到所有的偶然事件中，而不仅仅是那些在均衡博弈中发生的事件**。例如，子博弈精炼纳什均衡要求在绑架事件中薇薇卡的行动是最佳的，尽管盖伊的策略（不绑架 / 撕票 / 释放）将导致均衡博弈中不出现绑架事件。

定义 8.1

为完全信息博弈制定的子博弈精炼纳什均衡（Subgame Perfect Nash Equilibrium for Game of Perfect Information）：对于完全信息博弈而言，在每一个决策节点处，源于这一策略的行动能够使参与者的收益最大化，那么这个策略就是子博弈精炼纳什均衡。

下一章我们将给出一个关于子博弈精炼纳什均衡的更为正式的定义，它包含了完全信息博弈和不完全信息博弈这两类博弈。我们将重点讨论完全信息博弈。我们的方法就是当假定其他参与者的策略最优时为每一个参与者制定出详细的策略。我们提到过一个为每一个参与者的信息集指派一项行动的策略。盖伊的策略模板如下所示：

> 最初的节点处，______ [填入绑架或者不绑架]。
>
> 如果绑架事件发生且付赎金，那么 ______ [填入撕票或者释放]。
>
> 如果绑架事件发生但没有付赎金，那么 ______ [填入撕票或者释放]。

一个特别的策略会填入上面的三个空白处。薇薇卡的策略模板为：

> 如果绑架事件发生，那么 ______ [填入付赎金或者不付赎金]。

逆向归纳法（backward induction）包括从博弈中最后一个决策节点入手填入每一个参与者的策略模板，从而得出最佳行动方案，接着继续分析直到表中所示树状图的第一个节点。我们以这种方式得出子博弈精炼纳什均衡。通过案例了解这种方法是最简便的途径。

对于绑架博弈而言，最终两个决策节点是与盖伊是撕票还是释放奥兰多的决定相联系的。在每个节点选择盖伊的最佳行动。在薇薇卡选择付赎金的节点，盖伊的最佳行动是释放人质；而在薇薇卡选择不付赎金的节点，

盖伊的最佳策略是撕票。我们已经得到在这两个决策节点处盖伊的行动方案，可以在上面的两个空白处填入盖伊的策略模板。

对于这两个决策节点而言，下一步就是用与盖伊的最佳行动相联系的收益取代这个树状图（其中的任何一个案例都由撕票和释放人质这些分支组成）。这个取代过程在图 8—3 中有详细的解说。其中的假设就是，当我们面对这些决策节点时，盖伊将会选择最佳行动，所以我们很可能用与盖伊做出的选择相关的策略取代这些分支。

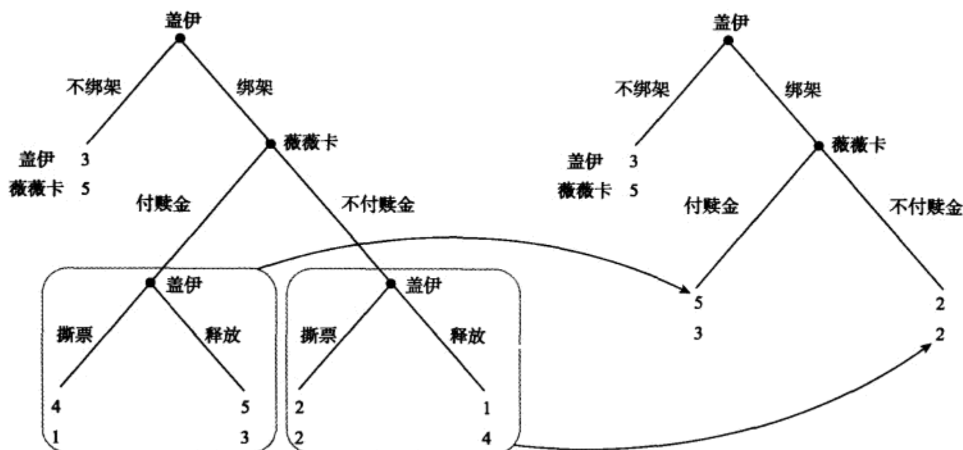


图 8—3　具体的转化步骤

在分析薇薇卡的决定时，我们将重复这一步骤。我们将对她付赎金与不付赎金这两种情况产生的收益进行比较。我们得出的结果是薇薇卡的最佳行动是付赎金。因此，我们找到了薇薇卡在决策节点处所做的最佳行动，从而得出关于她的子博弈精炼纳什均衡。我们用既定行动方案产生的收益取代产生薇薇卡决策点的这个树状图，从而得到了图 8—4 所示的博弈。这个博弈仅剩下一个决策节点，这个节点就是盖伊是否会有绑架行为。盖伊的最佳行动就是实施绑架。

正如刚刚所述，子博弈精炼纳什均衡（在这场博弈中这是唯一的）让盖伊在最初的节点处选择绑架，如果盖伊得到薇薇卡付赎金的消息，他将选择释放人质，而如果薇薇卡不付赎金的消息让盖伊得知，他会选择撕票。对薇薇卡而言，子博弈精炼纳什均衡就是付赎金。从而，唯一的一个子博弈精炼纳什均衡就是（绑架 / 释放 / 撕票，支付赎金），这个就是我们之前讨论过的有说服力的纳什均衡——唯一一个不依赖于薇薇卡不付赎金这一不可置信的威胁的均衡。



图 8—4　通过逆向归纳法得到盖伊的最初选择

我们刚刚理顺了一下逆向归纳法的解决方案。我们用广义的形式进行归纳，借助诱导推理的方式找到结果。

完全信息博弈的逆向归纳运法法则：

1. 为每一个最后的决策节点找到最佳行动。
2. 在每个这样的决策节点，用有相应收益的抉择点取代树状图的一部分，设想出一个最佳博弈。

PDG

3. 重复前两个步骤直到最初的决策节点。

子博弈精炼纳什均衡中最明确的就是这个概念，即参与者预料到其他参与者随后会做出最佳行动。例如，当我们用图 8—3 中的博弈得出薇薇卡的最佳行动时，我们假设她确实预料到了不管她做什么，盖伊会做出最好的回应。确切地说，当她考虑付赎金时，她预期盖伊释放奥兰多；当她不打算付赎金时，她预期奥兰多会被杀。在两种情况时，这对盖伊而言是最佳选择。因此，通过设想参与者的最佳策略打消对这个不可置信的威胁的顾虑，不管在这场博弈中他们充当什么角色。

对于一个完全信息博弈而言，这个步骤通常会产生一个问题。在每一个决策节点处通常会有一个最佳行动，这意味着你能够经常确认你要用来取代树状图的那一部分的收益。因为没有理由可以让这个步骤发挥作用，你可以回到最初的节点处去求解这个博弈。换一种方式，逆向归纳法包含了所有的子博弈精炼纳什均衡，而且这里至少存在一个子博弈精炼纳什均衡。



点拨

子博弈精炼纳什均衡的存在：在完全信息博弈中，至少存在一个子博弈精炼纳什均衡。

尽管对一个完全信息博弈而言，子博弈精炼纳什均衡的存在是确定无疑的，均衡的唯一性却得不到保障：在决策节点处没有单一的最佳行动时，子博弈精炼纳什均衡就会具有多重性。在种族歧视和运动队博弈中都出现了这种情形。



点拨

每一个子博弈精炼纳什均衡都是一个纳什均衡，但并非每一个纳什均衡都是一个子博弈精炼纳什均衡。

8.1 理解测验

用图8—5所示的博弈求解子博弈精炼纳什均衡

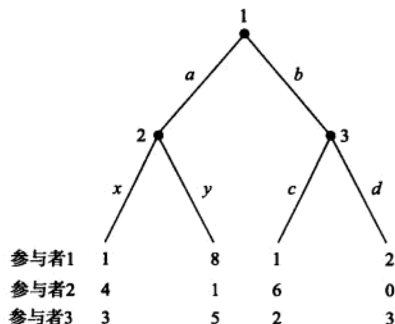


图 8—5

这里有另外一种说法，子博弈精炼纳什均衡比纳什均衡的要求更高。纳什均衡要求对那些达到均衡条件的决策节点而言，规定的行为是最佳的，但是子博弈精炼纳什均衡还有附加条件，那就是对那些尚未涉及的节点而言，所规定的行为也是最佳的。举例而言，（绑架 / 释放 / 撕票，付赎金）既是一个子博弈精炼纳什均衡，也是一个纳什均衡，而其他的四个纳什均衡不是子博弈精炼纳什均衡。

为了帮助我们更加顺利地使用这个逆向归纳法，我们列举了一些存在不当行为的案例。第一个例子，古巴导弹危机，苏联的不当行为是在古巴放置核导弹，问题是美国如何在避免发生核战争的前提下清除该地区的核武器。第二个案例，我们讨论了公诉人与安然公司高管订立协议的战略。最后一个案例，精明的种族中立主义者针对种族歧视问题制定了许多战略，当一个棒球经理人开始雇用非职业运动员时，这个棒球经理人的种族歧视行为将受到既定战略的惩罚。

案例研究

古巴导弹危机

1962 年 10 月 16 日上午 10 点召开参谋长联席会议（JCS）：

会议主席说他将于 11 点 45 分会见总统。

麦基将军（Mckee）：一旦允许一些场所进行导弹研究，卡斯特罗就可以对美国的任何攻击行为进行恐吓报复，他可以推迟自己的报复计划直到所研制的导弹足以发动一场核战争。

肖普将军（Shoup）：苏联可能试图在避免给自身招致核危机的前提下对美国进行核威胁。JCS 认为核威胁是相当严重的，所以需要美国采取军事行动消除这种威胁。

惠勒将军（Wheeler）：喜欢先进行航空偷袭，接着入侵他国。

麦基将军综合分析空袭和海上封锁的高效运作预测避免遭袭击的概率。

1962 年 10 月 14 日，美国确认苏联核导弹出现在古巴。当时正值冷战时期，苏联和美国作为世界上两个超级阵营处于对峙状态，双方都有足够实力发动一场具有空前规模的破坏性战争。苏联将这些武器推进到距离美国海岸线不足 100 米的地方。图 8—6 给出了 SS–4 (MRBMs) 和 SS–5 (IRBMs) 的射程范围。美国总统肯尼迪及其顾问面临的问题是如何在这些核武器投入使用之前将其撤离古巴，美国中央情报局估计需要 10 天时间。



图 8—6　苏联放在古巴的核导弹射程

美国和苏联领导人所面临的一系列决策如图 8—7 所示。美国首先要决定是封锁岛屿以阻止任何其他苏联船只到达古巴，还是在这些导弹投入使用之前发动空袭摧毁它们。如果采纳后一策略，决策结束，是否封锁岛屿的决定交给了苏联，苏联必须决定是保留导弹还是撤走导弹。如果他选择后者，美国又要决定是否发动空袭。

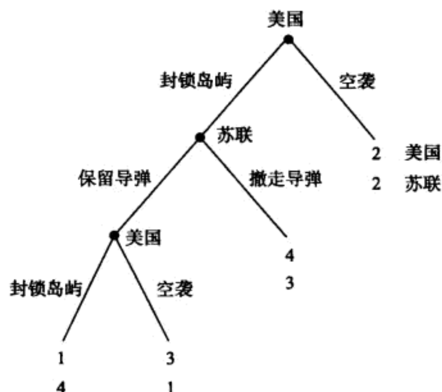


图 8—7　古巴导弹危机

反映到收益上，美国希望不发动空袭苏联就能撤走导弹，因为空袭有可能引发更大范围的战争。然而美国更希望摧毁这些核武器而不是让它们继续存在。苏联最希望保留核武器，但是也希望避免遭到空袭。

我们将使用逆向归纳法找到一个子博弈精炼纳什均衡。我们从最后一个美国决策节点开始，当美国封锁岛屿且苏联保留核武器时就到了这个节点。此时是美国发动空袭的最佳时机。图 8—8 中又具体描述，有一个箭头指向了空袭这个分支。你瞧，如果到了美国的第二个决策节点，最终的收益为（3,1），因为美国的最佳策略是发动一场空袭。

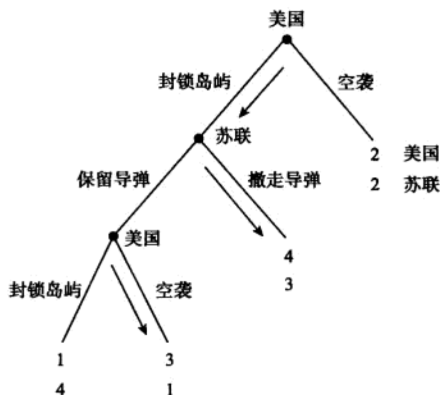


图 8—8　针对古巴导弹危机的子博弈精炼纳什均衡

向上分析这个树状图，到了苏联的决策节点：或者是撤走导弹得到3个单位的收益或者是保留导弹得到1个单位的收益，因为苏联预料美国将对空袭做出回应。鉴于撤出导弹能赢得更高的收益，箭头指向了这个行动。最终我们到了最初的节点处，如箭头所指，作为回应，苏联将撤走导弹。美国封锁岛屿的收益为4，空袭的收益是2，所以美国的最佳行动是封锁岛屿。

接着这里会有唯一的一个子博弈精炼纳什均衡，美国的策略是封锁岛屿，并且如果苏联不撤走导弹，美国将发动空袭摧毁这些导弹。但如果美国封锁古巴，苏联的最佳策略是撤走这些导弹。

事实上，美国确实组织海军进行了封锁，而苏联撤走了导弹。但是危机期间发生的不止这些。1962年10月26日，当时的苏联领导人赫鲁晓夫发表了一封信，建议在美国承诺不侵犯古巴的前提下撤走导弹。紧接着在第二天，赫鲁晓夫又提出了新要求，即要求美国也从土耳其撤走导弹，这个案例与古巴的案例很类似，都与苏联关系紧密。肯尼迪的兄弟、美国代理将军罗伯特•肯尼迪提出了一个不光彩的协议，不理睬赫鲁晓夫的第二个要求，但是告诉他美国政府同意他提出的不侵犯古巴的要求。

安然公司与检查特权

肯尼思•莱（Kenneth Lay）创建于1985年的安然公司被福布斯誉为世界上最具创新精神的公司之一。2000年该公司以超过10 000亿美元的收入在世界500强中位居第七。接着有新闻报道称这家公司存在财政问题，有报道称安然用金融贷款维持着虚假的繁荣，混淆视听。2000年8月，安然公司的股票是90美元每股，到2001年底跌到了1美元每股，致使这家公司面临破产危机。接踵而至的是各方的控告起诉，该公司的许多高级管理人员，包括CEO杰弗里•斯基宁（Jeffrey Skilling）和莱或者为自己的罪责服刑或者被判有罪。然而还未等到最终审判结果，莱就死于心肌梗塞，终年64岁。在白领犯罪的历史上，安然公司还残留着一件著名的案例。

在处理这类复杂的案件时，原告采取了一个策略：分化公司领导阶层。在与高级管理人员交换证词时，牵涉的官员将被“分化”。例如，大卫•迪莱尼（David Delainey）是一个中层行政人员，他被控有罪，作为被释放的条件，他需要跟原告合作。他根据传言提供了关于CFO安德鲁•法斯多（Andrew Fastow）的一些信息，这个CFO策划了公司的许多财政案件。接着法斯多被判有罪，处以十年监禁。他提供了斯基宁的犯罪证据，避免了被判处更长时间的监禁。作为公司领导层的最高人物，没有更重要的人物可供斯基宁和莱揭发了。

这个广义形式的案例如图8—9所示，它的参与者包括公诉人、迪莱尼和法斯多。公诉人首先要决定是跟迪莱尼签订协议，还是与法斯多订立协议，抑或是不签订任何协议。如果她决定订立一个协议，那么这个协议所针对的对象要决定是否接受协议。如果协议与法斯多相关，那么博弈结束。如果协议与迪莱尼相关且他拒绝接受，那么博弈结束（因为公诉人只是在考验他们）。如果迪莱尼接受了协议，公诉人要决定是否向法斯多提出一个协议。要记住在与法斯多的对决中，公诉人处于更有利的地位，因为他掌握着迪莱尼的证词。这场博弈将以法斯多接受或者拒绝协议结束。这场博弈的收益取决于公诉人是否在审理高管人员时附加更多的重要条件，从而决定是否根据申请予以减刑。这些行政工作人员最大的愿望就是将刑期减到最短，因此他会向公诉人提供他所掌握的所有证据。换句话说“小偷中间没有英雄”。

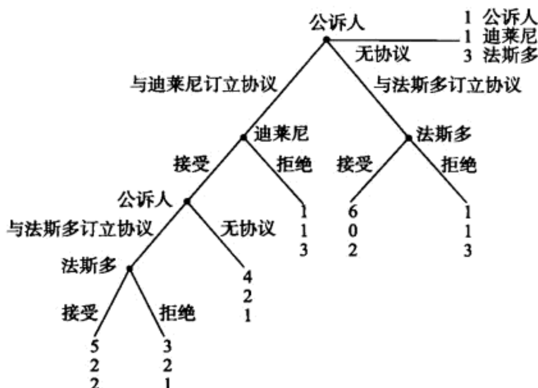


图8—9 安然公司的案件

这场博弈可以使用逆向归纳法进行分析，具体如下（因为我们得到了最佳行动，你可能发现采用图8—8的方法，在图8—9中附加箭头是很有帮助的）：

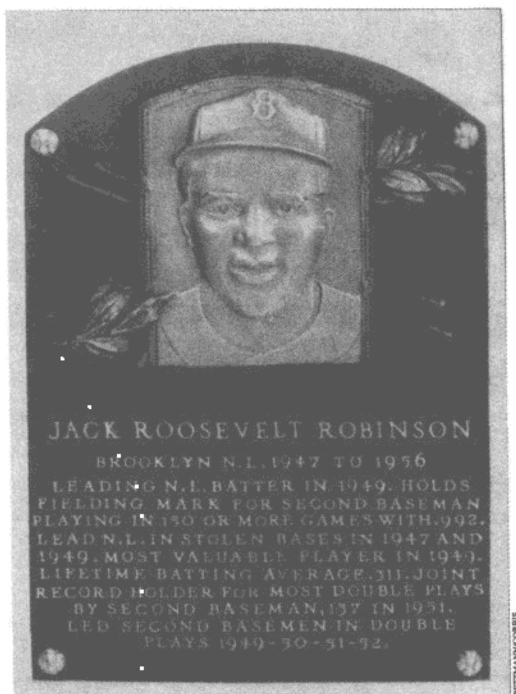
1. 最后的决策节点，法斯多要决定是否接受分化协议。如果他接受协议所得收益为2，拒绝接受所得收益为1，鉴于这一事实，对他而言，接受分化协议是最佳选择。

2. 继续分析树状图直到顶端，我们发现公诉人的决定是是否向法斯多提出协议（假设她已经向迪莱尼提出了协议并被接受）。她提出协议所得的收益为 5（我们假设法斯多的最佳选择是接受协议），而如果没有协议，她所得的收益为 4。这样一来，作为证据的交换条件，她会同法斯多订立协议。
3. 在决策节点处，迪莱尼面对一个协议，他会明智的选择接受，因为接受协议会得到收益 2，而拒绝接受所得收益仅为 1。我们注意到不管情况如何变化，迪莱尼所得的收益是完全相同的，因为他关心的只是自己的刑期而不是法斯多是否接受协议。
4. 在决策节点处，法斯多首先面对一个协议，他明智的拒绝了。这样做产生的收益为 3，而接受协议产生的收益仅为 2。（如果不是迪莱尼已经接受了分化，法斯多根本不担心自己的案件。）
5. 在最初的决策节点，公诉人面临三种选择，或者向迪莱尼提出协议（收益为 5），或者向法斯多提出协议（收益为 1，因为这个协议会遭到法斯多的拒绝），或者不提出任何协议（收益为 1）。她会向迪莱尼提出协议，毫无疑问，迪莱尼会接受协议，这将导致法斯多接受协议。

在这场与子博弈精炼纳什均衡相关的博弈中，公诉人会向迪莱尼提出协议，而他会接受，公诉人也会向法斯多提出协议，他也会接受。这样，这个案例就是控告斯基宁和莱。

种族歧视及运动队

1947 年布鲁克林道奇队（Dodgers）招入杰基•罗宾森（Jackie Robinson）之前，美国职业棒球大联盟中从来没有过黑人。实际上在分队比赛时，具有棒球天分的黑人球员会被安排到黑人俱乐部。打破种族障碍让道奇队拥有了更多有才华的球员——先后有黑人棒球运动员罗伊•坎帕内拉（Roy Campanella），乔•布莱克（Joe Black），唐•纽康（Don Newcombe）加入，这诱使其他俱乐部采取类似的措施，招募被忽视的具有天赋的运动员。1959 年，最后一个专业棒球俱乐部——波士顿红袜队让帕姆普西•格林（Pumpsie Green）站上二垒打完比赛后也进行了整合。



杰基•罗宾森纪念大厅的功名牌匾

为了探讨种族歧视的概念，我们假设一种情形，所有的专业运动队每年举行一次：应征的是非专业运动员，特别是刚刚从高校毕业的。假设有两支球队，道奇队和红袜队，有四名棒球运动员应征，编号为 1,2,3,4。棒球运动员按照种族和才能（概括为比赛技能）划分，如表 8—1 所示。比较才能，1 号棒球运动员优于 2 号，2 号优于 3 号，3 号优于 4 号。1 号和 4 号球员是黑人，2 号和 3 号球员是白人。每个球队都乐于征募球员以最大化球队的整体价值。道奇队不考虑种族，他们认为较有才能的球员具有更高的价值而不论种族。作为对照，我们假设红袜队既看重种族也在乎才能。他们认为这两个最好的白人球员具有最高价值。

表 8—1　棒球运动员的技能和种族

球员	才能	种族	道奇队的收益	红袜队的收益
1	30	黑	30	20
2	25	白	25	25
3	22	黑	22	22
4	20	白	20	10

球员征募让两个球队开始选择球员，图 8—10 所述的是棒球队征募的扩展式博弈。今年道奇队得到优先选择权，能够从四个球员中任意挑选。红袜队在剩余的三个球员中选择一个。然后在剩余的两个中，道奇队再选择一个，最后一个归红袜队。

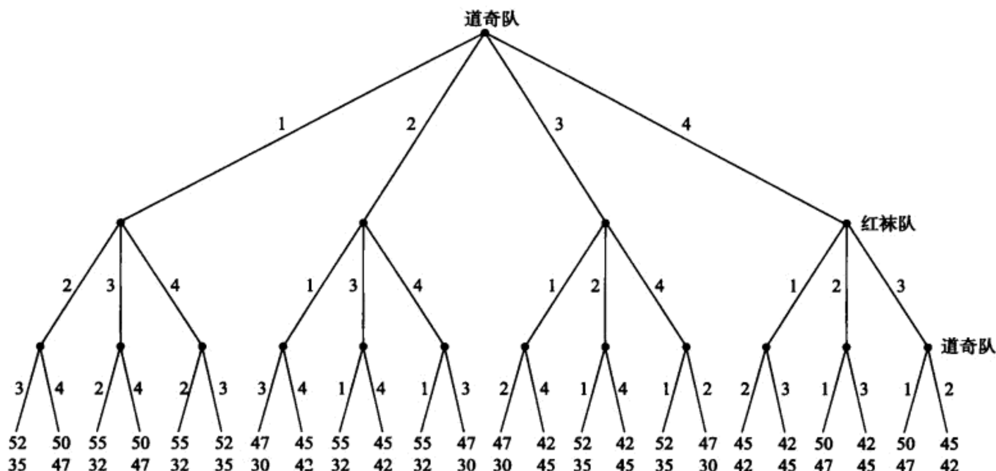


图 8—10　棒球队招募

尽管比起本章前面介绍的博弈，扩展式博弈有更多的分支，但是逆向归纳法同样适用。考虑一下与道奇队第二次选择相关的最后 12 个决策节点。征募过程中最佳博弈就是在剩余的运动员中挑选最优的。（然而正如我们所看到的，征募过程中这个要求并不很及时。）第一轮选择，道奇队选择了 1 号球员，红袜队选择了 2 号球员，接着道奇队可以在 3 号和 4 号球员中选择。他们会选择 3 号，也就是说他们最后选择的球员是 1 号和 3 号，收益为 52（两个球员的总价值），而红袜队选择了 2 号和 4 号球员，收益为 35。接着我们用图 8—11a 所示的收益（52,35）代替树状图的那一部分。在道奇队剩余的其他 11 个决策节点重复这个步骤。

假设存在图 8—11a 所示的博弈，红袜队有四个决策节点，我们需要为这四个节点寻找最佳行动。与道奇队相关的决策节点在第一轮选择了 2 号球员。接着红袜队要在 3 号和 4 号球员中选择，两种选择都产生 32 个单位的收益。如果他们选择 3 号球员，那么在下一轮道奇队选择 1 号球员，这样红袜队最终选择的是 3 号和 4 号球员。如果红袜队先选择了 4 号，那么在下一轮选择中，道奇队选择 1 号，因此红袜队最终拥有 3 号和 4 号球员。这两种选择不会影响到球队的收益，道奇队得到的收益是 55，红袜队得到的收益是 32。对其他的三个决策节点进行相同的分析，我们得到图 8—11b 所示的博弈。

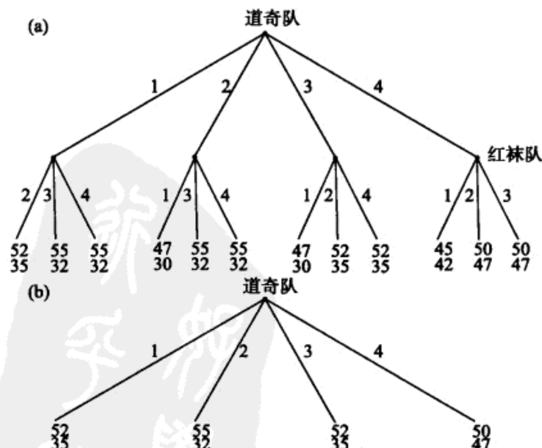


图 8—11　逆向归纳法在棒球队招募博弈中的运用

我们的博弈分析已经到了道奇队需要作出最初的选择的这个节点。如图 8—11（b）所示，他们会选择 1 号棒球运动员，基于此，红袜队会在剩下的球员中做出最佳选择，选择 2 号（最有潜力的白人球员），之后道奇队选择 3 号。具有优先选择权的道奇队最初选择了最有才能的球员（1 号），因此最终得到的球员是 1 号和 3 号，所得收益为 52。如果他们最先选择的是 2 号球员，那么紧接着红袜队会选择 3 号，道奇队在第二轮选择 1 号。就此情形，道奇队最终得到两个最具潜能的运动员，获得较高收益 55。如果道奇队在第一轮选择了 3 号球员，他所得收益为 52，如果选择的是 4 号球员，所得收益为 50。因此，对道奇队而言，最佳选择不是在第一轮选择最具潜力的球员，而是选择最具潜力的白人球员。

尽管看起来道奇队似乎有种族歧视观念，因为他们选择了能力较差的白人球员，但实际上他们没有种族歧视观念。实际上，红袜队是存在种族歧视的，这一点在其策略中被巧妙的掩饰了，道奇队却利用这种歧视赢得了优势。道奇队在第一轮招募时选择了最好的白人球员，因为他知道尽管如此他依然能得到最具才能的球员，只因这个球员是黑人，而注重种族的红袜队不会选择这个球员，从而道奇队得到了两个最好的球员。所以，道奇队利用了红袜队的种族歧视观念，组建了一支强悍的球队。打破种族歧视或许是最好的方法，而保持种族歧视会让参与者的利益受损。

值得注意的是这里有许多子博弈精炼纳什均衡。记录下整个策略内容（道奇队的 13 元组和红袜队的 4 元组），我们可以看到这个多样性源于完全信息博弈下唯一最佳行动的存在。例如，这里有一个子博弈精炼纳什均衡，红袜队会在道奇队选择 2 号球员后选择 3 号球员，另外一个子博弈精炼纳什均衡让红袜队在道奇队选择 2 号球员后选择 4 号球员。然而所有的子博弈精炼纳什均衡都会让道奇队最后拥有 1 号和 2 号球员。

棒球队的种族障碍被打破之后，种族歧视在球队的决策中消失了吗？实际上，如果球队的经理人没有种族歧视，我们希望白人和黑人球员能有相同的表现。如果经理人对黑人运动员存在偏见，只有在一个黑人运动员确实比白人球员优秀很多的情况下，球队经理人才会选择黑人球员。据此我们得出一个结论，种族歧视的标志在于黑人球员优于白人球员。

1947 到 1975 年的证据表明，一般而言，黑人球员在许多方面的表现优于白人。表 8—2 给出了对博弈三个阶段的统计。如果你并不熟悉棒球，你只需记住任何项目中较高的数字就表示较好的表现。19 世纪 70 年代早期，在杰基•罗宾森到来 25 年后，黑人击球手普遍有较高的击中率，更大力的击球和更多的盗垒。评价投球手表现的方法告诉我们一个类似的道理。尽管对这种表现或许存在另外一种解释，但是棒球球队坚持种族歧视的前提是一定要保证的。

表 8—2　棒球球队的种族及表现情况

年份	种族	平均长击率	平均长击率	安打	全垒打	盗垒
1947–1960	白	0.261	0.393	144	13	4
	黑	0.280	0.455	154	20	10
1961–1968	白	0.251	0.380	138	14	4
	黑	0.269	0.421	148	17	14
1969–1975	白	0.251	0.369	138	12	5
	黑	0.272	0.419	149	16	16

8.2 理解测验

分析图8—12中的博弈，找到子博弈精炼纳什均衡。

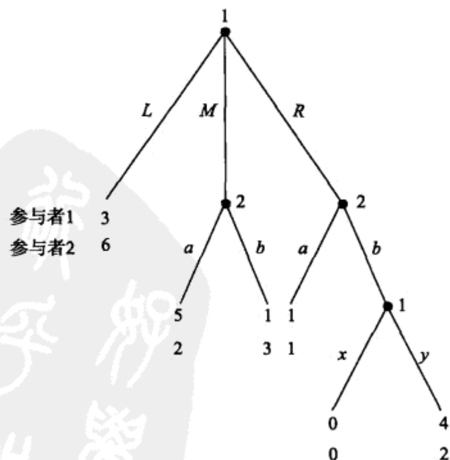


图 8—12

8.2 等待博弈：抢先博弈与消耗战

大多数美国人一生中有两到三年的时间是在排队等待。

《哈利波特与混血王子》(*Harry Potter and the Half-Blood Prince*) 这部电影在你居住的城市上映，你正在考虑什么时候去剧院。你希望得到票并且是好座位，但是你不愿排队等候。什么时候去买票合适呢？现在假设你正在剧院重新排队，希望得到贵宾区的座位。电影就要开始了，你会继续等待吗，你希望排在你前面的人会放弃排队，或者回到你原来那个没有爆米花和饮料的座位？现在我们将场景设定在飞机场，你在登机口等待登上西北航空的班机。因为西北航空提供开放式座位（“先到先得”），你正在犹豫是加入排队的行列还是继续舒服的坐着。现在假设你正在家里上网，你看到了 eBay 的一项拍卖活动。这次拍卖在一小时内结束，你正在决定是现在投标还是稍作等待接近结束时再投标。

在与时机有关的决定中，所有的情形都是有共同点的。你什么时候排队？什么时候放弃排队？什么时候投标？这些都有固定的策略，因为你的决定源于你所认为的其他参与者的行动。如果你认为会有很多哈利波特影迷前往剧院，那么你会选择尽早赶去买票。如果排在你前面的人正在聊他们错过了午饭，你会选择放弃排队回到你原来的座位上。如果你环顾西北航空登机口四周，发现许多人都在轻松地交谈，你会愿意等会儿再去排对。

本章我们讨论参与者关于时机决定的策略。这些情形包括先于其他人行动是否有利（正如在电影院排队这个案例），还是应该晚于其他人（正如在 eBay 竞标的案例）。第一种情形被称为**抢先博弈**（Preemption game），目的在于试图超越他人。具体来说，抢先博弈是这样的博弈，每个参与者要决定何时采取行动，当下列情形发生时参与者获得较高收益：（1）先于其他人行动；（2）行动之前等待更长的时间。比如，在西北航空的开放式座位上，参与者愿意坐尽可能长的时间，然后排到队伍最前方。一个参与者犹豫的时间越长，其他人越有可能超越他先行动。所以我们会看到均衡博弈包含参与者缺乏耐心而采取行动，他们在行动之前没有等足够的时间。

作为对照，**消耗战**（war of attrition）是时机博弈，参与者在以下情况下能获得较高的收益：（1）其他参与者较早行动；（2）自己较早行动。特别的，如果一个参与者想要行动，他希望自己尽早行动。在所允许的范围内，你希望其他参与者超越底线首先行动，但是如果是你打算超越底线，你愿意现在就行动而不是等待以后。一个人等待其他人采取行动的时间越长，损失的收益越大。均衡博弈要求参与者有足够的耐心，在行动之前等待足够长的时间。

8.2.1 抢先博弈

我们考虑一下关于西北航空开放式座位的情景。尽管登机口的排队等待博弈涉及许多乘客，我们为简化博弈，设定只有两个乘客，决定是坐在座位上还是排队。假设只要有一个乘客排队其他人就会跟随其后。（她先于其他许多的乘客，而这些乘客是我们不曾设定的。）排在队伍最前面的收益为 30，第二位为 20。与排队等候相关的成本如表 8—3 所示。等待的时间越长，成本越高。也就是说，另外一个参与者所花费时间的数值越大，等待的时间越长。例如，如果一个参与者已经等待了 1 个时间单位，第二个时间单位的成本为 7（=12−5），如果他已经等了两个时间单位，第三个单位的成本为 9（=21−12）。乘客的收益就是他们在队伍中所排的位次减去排队所花费的时间。

表 8—3　排队等待的成本

花在排队上的时间单位	时间单位的成本
1	5
2	12
3	21
4	32
5	45

广义形式如图 8—13 所示，乘客 1 首先决定是行动（比如排队）还是等待。如果他行动了，那么博弈结束，他所得收益为−15（排在队伍最前边所得的收益 30 减去等待所花费的成本 45），乘客 2 所得收益为−25（收益 20 减去排队花费的成本 45）。如果乘客 1 等待，那么乘客 2 决定是否行动。如果他等待就轮到乘客 1 行动，以此类推。这种情形最多会持续 5 个回合。如果在

最后的决策点乘客 2 决定等待，我们就会猜想队伍最前方的参与者是随机决定的，在此情形下每个乘客的收益都是相同的 25= (1/2) ×30+ (1/2) ×20（此时不存在等待成本）。

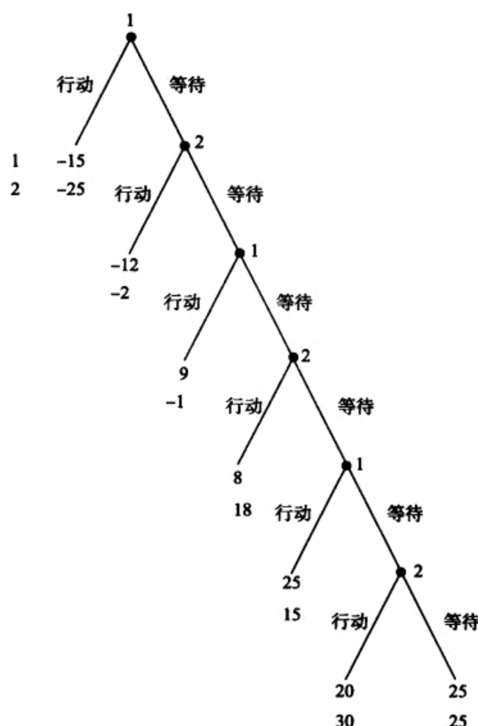


图 8—13　排队博弈

现在我们用逆向归纳法分析这个博弈。如果每个参与者都放弃排队，那么乘客 2 处于他最后的决策点，此时他会快速跑到队伍的最前方得到收益 30，而不是继续等待那 50% 排到最前方的概率（收益为 25）。得益于这个首先发生的行为，在她的最后决策点乘客 1 会排队，这样产生的收益是 25，相对于等待到最后以排在第二位结束（收益为 20），这样做是较明智的。通过逆向归纳法分析，乘客 1 认为如果她等待，乘客 2 会跳出来排队。返回到乘客 2 的第二个决策点，他会选择行动并因此得到收益 18 而不是等待，如此一来，乘客 1 在下一轮跑到了他的前面。在乘客 1 的第二个决策点，他可以排队得到收益 9 或者等待得到收益 8，所以他选择行动。在乘客 2 的第一个决策点，他会果断地选择等待得到 -1 的收益，而不是排队的收益-2。阻止他进入队伍的附加成本是相当巨大的，尽管他知道乘客 1 将在队伍中排在第一位，这个附加成本是四个单位而不是三个单位的在队伍中等待的成本。在乘客 1 最初的决策点，她很清楚她需要等待。不管她是进到队伍中还是等待，她最终都会排到最前面，因为乘客 2 在下一轮会选择等待。因此，她更愿意等待以减少排队消耗的总时间。对乘客而言，唯一的子博弈精炼纳什均衡就是在开始的两轮选择等待，在第三轮和乘客 1 一起开始排队。

之前的方案给出了抢先博弈中均衡博弈的大体特性：参与者以相当快的速度采取行动。在均衡博弈中，乘客 1 的收益为 9，乘客 2 的收益为-1。将这些收益与他们等待 5 个回合所得到的收益与乘客 2 排在队伍最前方得到的收益进行对比。他最好以得到 30 的收益结束，但是乘客 1 会得到 20 的收益，因为她避免了排队等待。这种低效率的博弈产生了 —— 没有等待足够长的时间，因为每个乘客都存在急切心理，这导致他们过早的排成一队。

关于抢先的影响的一个令人震惊的案例发生在医生和居民的交易市场上。在 20 世纪初医院雇用医学院的毕业生作为医生，这样既给医院带来了廉价劳动力也给学生提供了锻炼机会。医院经常与其他医院竞争以得到最好的学生。

> 解释这种竞争的一种方式，是医院试图设定一个日期，他们会先于自己的主要竞争对手，与医学院共同制定一项可执行的协议。结果是，他们在这一天定下了许多主要协议，从在医学院学习的最后一个学年开始缓慢向前分析。医院认为这种方法代价昂贵，效率低，医院必须在不知晓学生毕业成绩和在校表现的前提下进行任命，并且学生和医学院发现，寻找优秀学生的行动打乱了许多学年的正常教学计划。

在 19 世纪 40 年代初期，医院的竞争从大学三年级就开始了。在医学院的学生成为医生之前，医院会提前与之签订两年的劳动合同。直到所有医院一致同意制定一个共同的日程表，这种低效率的竞争才得以终止，这时已到了 19 世纪 40 年代后期。

8.2.2 消耗战

消耗战源于一种军事冲突类型。双方持续交火，每一方都希望对方能迅速放弃进攻或者投降，他们都希望能比对方坚持更长的时间。一个经典的案例就是第一次世界大战中的西部战线，敌对双方军队都躲在堑壕里，德国军队占据一方，英美军队占据另一方。这个战区具有持久性的特点，战争持续了数月，伤亡的人数是固定的。每一方在等待对方投降的过程中都付出了惨重的代价。

图 8—14 描述了只有两个参与者的消耗战。自己首先行动的收益为 0，而让对方首先行动的收益为 100。如果双方都不行动，那么每一方的成本消耗都是 10。类似的，对这个抢先模型而言，如果到博弈的最后都没有人采取行动，那么此时的收益将受到哪方必须行动的随机决定的影响。

使用逆向归纳法，我们注意到每一个决策点处，子博弈精炼纳什均衡让参与者等待而不是行动。例如，在她的第二个决策点，参与者 1 料想参与者 2 会等待，所以她也选择了等待，这样产生的额外收益是 10，如果行动将损失的收益是 20。这个策略让所有的等待产生了更多的收益——100，这些收益被消耗在了一个决策点上，这个决策点的总收益仅为 20（每个参与者得到 10）。

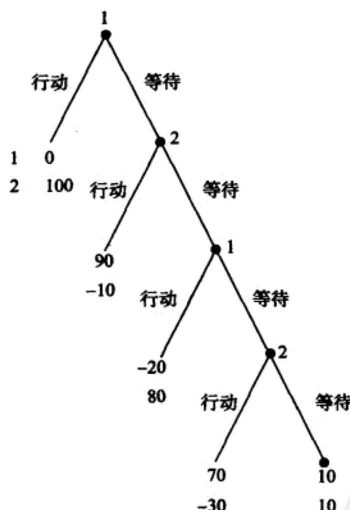


图 8—14　消耗战

这是消耗战均衡博弈的一个共同特性：参与者等待的时间过长，因为每一个参与者都在拖延等待的时间而不是缩减。对他们而言，一个更可行的结果是投硬币决定谁先行动，一旦做出决定，这个人会立刻采取行动。这个方法会给他们带去 $50=(\frac{1}{2})\times0+(\frac{1}{2})\times100$ 的期望收益，这个结果极大地超越了均衡收益 10，而两者的区别在于等待的成本。问题在于投硬币这个方法是不可行的。如果投硬币的结果是让参与者 1 行动，那么根据均衡博弈，参与者 1 会说这只是玩笑，而选择等待。

专业棒球俱乐部的经理人和参与者很多时候都在进行消耗战。没有成功签订协商一致的劳动合同，棒球运动员开始罢工，导致 1972 年取消了 86 场比赛，1981 年取消了 712 场比赛，1994—1995 年取消了 938 场比赛。尽管经济成本在不断增加，每一方还是希望对方能做出必要的让步，签订一个新的劳动合同。经理人和运动员都知道进行一场消耗战要付出昂贵的代价。

8.3 人类在理智地使用逆向归纳法吗

人类的行为是很难理解的。本书的任务在于说明博弈论如何帮助我们解决这个问题，但是我并不想骗你说这个问题已经解决了。社会科学家使用博弈论和其他工具已经使这个问题的研究有了一些进展，但是还有许多问题是我们不理解的。为了得到一个有助于解释这个疑问的方法，接下来我们将提供两个对逆向归纳法的评论。第一个评论就是，许多策略中，逆向归纳法所预言的并不能跟人类的行为很好的契合。8.3.1 小节提供了证明这个评论的事实依据。8.3.2 小节研究了逆向归纳法的逻辑基础，这导致了以逆向归纳法假设内容为中心的问题研究。

8.3.1 实验性证据与逆向归纳法

篮球比赛临近结束的几秒钟，你正在带球，你必须决定是投篮还是把球传给其他队友。如果你传球，这个接到球的队友也要决定是投篮还是传球。投篮之前传球的次数越多，越有可能找到善于投篮的球员，他就越有可能投中并赢得比赛。同样是这个案例，每个球员都希望自己能够投篮。

假设时间只允许在队员利本和拉里之间进行三次传球。图 8—15 描述了这个扩展式博弈。利本可以投篮得到收益 4 或者把球传给拉里。拉里或者投篮得到收益 6

或者再把球传回给利本。利本可以投篮得到收益 8 或者把球传给拉里，此时因为所剩时间有限，拉里将会投篮并得到收益 10。

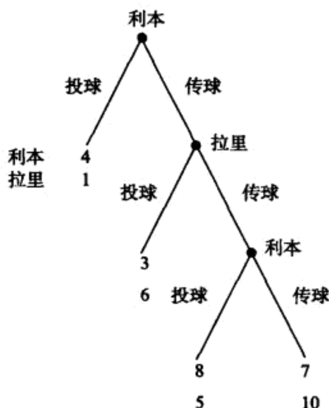


图 8—15　篮球比赛中最后一投

通过透析球队我们知道最好在投篮之前多传几次球，但是逆向归纳法给出了一个非常不同的预言。从最后的决策点开始，利本决定是投篮还是进行最后一次传球，将球传给拉里。因为投篮他可以得到收益 8，但是传球仅得到 7，所以他选择投篮。回到拉里的决策点，他可以投篮得到收益 6 或者传球得到收益 5。拉里知道如果他把球传给利本，利本会投篮。因此，一旦拉里得到机会，他希望自己投篮。在最初的决策点，如果利本投篮，他仅能得到收益 4，但是比起传给拉里得到收益 3，还是自己投篮更为明智（利本知道一旦拉里拿到球他就会投篮）。子博弈精炼纳什均衡让参与者在每一个决策点都选择投篮，这就意味着利本要立刻投篮而不是传球以等待一个更好的投篮机会。

篮球比赛是一个存在很多不确定性的博弈，就像蜈蚣一样。图 8—16 描绘出它的广义形式，我们看到最终图的形状就像蜈蚣一样。这场博弈以两堆钱开始。参与者 1 决定是抢占较多的那堆——此时参与者 2 得到较少的那堆，博弈结束，还是把较多的留下。如果较多的那部分被留下，就给两部分都增加更多的钱，现在轮到参与者 2 做出同样的选择：抢占较多的部分还是把它留下。如图 8—16 所示，最初较多的部分有 40 个硬币而较少的部分有 10 个硬币。如果较多的部分被留下，那么总钱数就会翻倍，所以在第二个决策点，较多的部分有 80 个硬币，而较少的部分有 20 个硬币。每一个较多的部分被留下，总钱数就会翻倍。如图 8—16 所示，这场博弈最多会持续 6 个回合。如果在最后的决策点，参与者 2 留下了较多的那部分，那么最后总钱数会翻倍，参与者 1 必然会得到较多的那部分钱（只跟时间有关，跟抢占成功的百分比无关）。

	1	2	1	2	1	2	
	留下	留下	留下	留下	留下	留下	25.60 6.40
	抢占 (1%)	抢占 (6%)	抢占 (21%)	抢占 (53%)	抢占 (73%)	抢占 (85%)	
参与者1	0.40	0.20	1.60	0.80	6.40	3.20	
参与者2	0.10	0.80	0.40	3.20	1.60	12.80	

图 8—16　蜈蚣博弈

假设一个参与者的收益是由他所得到的钱数决定的，我们将使用逆向归纳法去分析这个博弈。最后的决策点，参与者 2 会得到较多的那部分钱，因为他喜欢 12.8 美元胜过喜欢 6.4 美元。回到倒数第二个决策点，参与者或者抢占到较多的那部分钱得到 6.40 美元的收益，或者留下那部分钱得到 3.20 美元的收益。假设参与者 1 认为参与者 2 会采取最佳行动，在下一个决策点抢占较多的那部分钱，那么她会得到 3.20 美元的收益。继续进行逆向归纳，我们发现对一个参与者而言，一旦有机会就抢占较多的钱才是最佳策略。所以逆向归纳法暗示参与者 1 在博弈开始时抢占 40 个硬币，此时博弈结束。

逆向归纳法为蜈蚣博弈提供了一个精确的方法。唯一的问题在于人们不按照此方案行事——至少没有严格遵守这个方案。我们用真实的钱，真实的在校大学生，把图 8—16 所示的博弈作为研究课题进行了 281 次试验。实验到了第 i 个决策点，图 8—16 所列出的抢占成功的百分比是人们取走钱而不将其留下的概率。例如，在第一个决策点，281 次实验中参与者 1 只有两次取走了 40 个硬币，占实验的 1%。279 次实验到了第二个决策点，有 18 次参与者 2 抢占了较多的那部分钱，所占概率为 6%。

理论表明一旦有机会参与者就会抢占较多的那部分钱。这个预测表明，在第一个决策点参与者 1 会抢占 40 个硬币，但是在实验中，这种现象发生的概率仅为 1%。

随着博弈的进行，抢占较多部分钱币的概率有固定的增加，所以存在的决策点越少，我们所观察到的行为越接近于这个预言。在第五个决策点，参与者 1 抢占较多钱币的概率为 73%。但即使到了最后的决策点，当参与者 2 要在 12.80 美元和 6.40 美元之间选择时，实验表明他选择后者的概率仍为 15%。

我们是否可以得出这样的结论，人类在制定决策时不使用逆向归纳法？并不。逆向归纳法只是我们理论的一个组成部分，理论和实践的不一致性可能导致一些拙劣的参与者选择这部分。让我们讨论一下怀疑的一些可能性。但是在此之前我要问一下，你认为为什么人类的行为会有别于理论的预测？在继续下面的内容之前说出你的想法。

确实存在人们没有使用逆向归纳法或者人们在使用它时被局限住了的可能。很重要的一点是，当仅有两个决策点时，实验结果和预测相一致的概率为 73%，参与者 1 会抢占 6.40 美元，这相对于理论而言不是一个坏的表现。此时参与者 1 必须期待一种行为 —— 预测参与者 2 下一步的行动计划，以便制定她的最佳方案。当剩下三个决策点时，参与者 2 不得不有两种期待，这是更具挑战性的。在那个决策点，实验结果和理论一致的概率仅为 53%。剩余的决策点越多，人们的行为背离理论预测的概率越大。或许人们使用逆向归纳法的范围是有一定限制的。

这个理论的另一个值得重新评价的部分是收益。实验的实施者决定使用多少钱而不是实验参与者的收益。这个理论是基于参与者只关心他能得到多少钱这一假设的，但是如果参与者有不同的表现呢？假设参与者关心的不仅仅是钱呢？这一点告诉我们为什么一些参与者，虽然不是很多，在最后的决策点没有选择抢占那部分钱，尽管这意味着选择 6.40 美元而不是 12.80 美元。或者是他们做错了，或者是这些钱没有准确地反映到收益上。

还有一种观点认为参与者不仅在乎他们得到多少钱，同样关心其他参与者得到多少。这一部分人可能具有无私精神 —— 他们希望其他人得到更多的钱（或者他们会为自己利用了他人而感到难过），甚至会牵连到那些不具有无私精神的参与者。假设你不确定与你搭档的参与者是否有无私精神。如果她是无私的，在下一轮博弈中她不会抢占这笔钱，你当然不会在这一回合抢那笔钱，因为你想让它增值后再抢占。所以一个自私的逆向归纳法的使用者会在几个回合的博弈中选择留下较多的那部分钱，如果她不确定她自身或者跟她搭档的参与者是否是无私的。

8.3.2 关于逆向归纳法的一个合乎逻辑的表述

我们把实验所得到的事实依据放到一边，重新检验一下逆向归纳法讨论的逻辑。假设这场博弈是非常具体明确的，正如图 8—16 所描述的。那么根据逆向归纳法的讨论，一旦到了第五个决策点，参与者 1 会抢占 6.40 美元，因为如果他不这样做，参与者 2 会在下一轮抢占这笔钱，仅给参与者 1 留 3.20 美元。但是为什么参与者 1 希望参与者 2 得到这笔钱呢？因为他相信参与者 2 是理性的，而这正是一个理性参与者应该做出的选择。现在让我们回到第四个决策点。如果参与者 2 没有得到 3.20 美元，他希望参与者 1 在下一轮得到这笔钱，原因在于参与者 2 相信参与者 1 认为他是一个理智的参与者，所以参与者 2 相信参与者 1 认为他会在第六个决策点抢占这笔较多的钱。基于这种信念，参与者 2 希望参与者 1 在第五个决策点得到这笔钱，因为参与者 2 相信他是理智的，从而比起 3.20 美元，他更愿意得到 6.40 美元。参与者 2 会在第四个决策点行动并得到 3.20 美元。继续用这一逻辑进行分析，在参与者 1 和参与者 2 之间理智是一种共同知识，这一点暗示着参与者 1 会在博弈开始时抢占 40 个硬币。

现在出现了一个难题。参与者 1 是不是应该在这笔钱增加到 6.40 美元时就抢占它呢，一个关键的前提在于参与者 1 相信参与者 2 是理智的，从而如果条件允许这笔钱增加到 12.80 美元，参与者 2 将会抢占这笔较多的钱。但是如果理智是一种共同知识，那么这场博弈该如何进行到第四个决策点呢？理智是一种共同知识就暗示着这场博弈会在第一轮结束，不可能持续到第四轮。所以是不是这笔较多的钱超过 40 个硬币就意味着理智不是共同知识呢？如果确实如此，原因何在？而且我们要确认一点，参与者 1 相信参与者 2 是理智的假设是否成立？参与者 2 先错过了这笔钱，否则，参与者 1 不会面对这样的选择 —— 是否拿走 6.40 美元。所以如果是 12.80 美元，她可能再次放弃吗？如果她放弃了，当这笔钱为 6.40 美元时，参与者 1 会放弃，以使得他最后能得到 25.60 美元吗？

图8—17所表示的这个扩展式博弈的特性是很有说服力的。乔恩首先行动，如果他选择b，那么简得到行动机会，之后乔恩再次行动。为了得到她的最佳策略，简需要预测如果她选择x，乔恩会有何方案，如果她选择y，乔恩又会如何做。使用逆向归纳法，简预测如果她选择x，作为回应乔恩会选择d——这导致简的收益为5；如果简选择y，作为回应乔恩会选择c——这导致简的收益为3。所以如果在最初的决策点乔恩选择b，简会选择x。

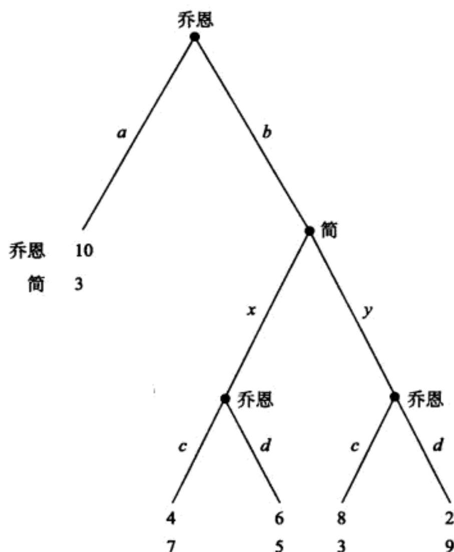


图8—17

我们注意到如果选择a，乔恩得到的收益为10，然而他选择b的收益不会超过8，实际上当所假设的博弈是最佳博弈时他所得的收益仅为6。换句话说，所有的策略让乔恩在最初的决策点选择a而不是b。从而，如果博弈到了简的决策点，她知道乔恩选择了一个严格占优的策略。此时，简是否会认为乔恩会用一个最佳的行动回应简的方案呢？

以简的决策点开始的有助于逆向归纳解决的一个讨论就是乔恩试图理智的行动，但是在最初的决策点他就很轻易地犯了一个错误：他原打算选择a，却错误地选择了b。如果我们认为他不会犯更多的错误，那么简得到的关于乔恩如何回应她的最佳答案就是他会采取最佳行动方案。但是或许他会犯更多的错误。乔恩可能受了药物的麻醉，所以他选择b，这样他可能会犯更多的错误。关于这个难题你是如何认为的呢？

本章总结与练习题

本章总结

本章首先指出纳什均衡的一个不足，并将之作为本章要解决的问题。在扩展式博弈中，纳什均衡仅需要参与者的策略针对那些既定信息确定一个最佳行动方案（或者是在完全信息博弈中，决策点），这是在均衡博弈的进程中所达到的。因为执行的行动是影响参与者收益的唯一因素，为一次偶然事件设计一个不合情理的行动方案是没有任何成本的。换句话说没有必要花费成本去制造一种假象。纳什均衡策略将为均衡博弈尚未涉及的决策点设计一个并不是最佳的行动方案。这个创造所担心的仅仅是令人不可置信的威胁，这个策略成为了一个纳什均衡。也就是说，仅仅因为一个参与者威胁说他要做什么，这个参与者的策略成了最佳策略，但是如果那个参与者被认为是虚张声势的，那么相信他的威胁将是不明智的。

这个不太现实的案例让我们考虑一个更加严谨的解决方案，使得我们能够用这个方案统治依赖于威胁的纳什均衡。对一个完全信息博弈而言，从均衡博弈所涉及的那些决策点到所有的决策点，子博弈精炼纳什均衡提出了一个完美的规则。逆向归纳法则为得出子博弈精炼纳什均衡提供了一种方法，而这种方法是通过这样的途径得来的，即解决决策点的完美的行为进而得出参与者在树状图中的运动轨迹，所有的方法都预测后一个博弈是最佳的。在有着透明信息的博弈中，这种运算法至少产生了一个子博弈精炼纳什均衡。

关于序列博弈的整体研究包括参与者面对一系列决策时的时机选择，这个决定要求参与者考虑是行动还是等待。所举的案例包括何时在登机口排队，劳工协会何时该举行一次罢工。在抢先博弈中，先于其他参与者行动是很有益处的，但是也存在这样的担忧：过早采取行动要付出高昂的代价。参与者最好的结果就是等待，然后成为第一个行动的。每个参与者不得不超越他人的诱惑产生了一个收益占主导的决定，参与者过早采取行动。作为对照，消耗战一般更有利于后一个采取行动的。但是在等待的过程中，这个参与者也要付出一定的代价。此时，这个参与者最理想的结果就是其他人立即行动。等待的诱惑源于这样的希望，其他参与者会采取行动以导致所有的参与者等待过长的时间。

尽管逆向归纳法对产生一个看似并不合理的纳什均衡很具有吸引力，但是我们已经使用某种方法和子博弈精炼纳什均衡使两种评论达到了同一水平。首先，许多实验说明（比如蜈蚣博弈涉及的主要实验），参与者的行动与均衡原理所预测的结果截然相反。尽管理论和现实之间所欠缺的讨论可能是因为这个模型的其他因素（比如我们所设想的参与者的行为）而产生的，可能存在这样的情形：他们合乎常理的行动并不完全符合逆向归纳法。其次，逆向归纳法存在逻辑上的矛盾。在运用这个概念的过程中，你努力进行着一个想法实验，当到达博弈的一个决策点时你主要思考什么是参与者的最佳策略。在制定最佳策略时我们假定其他参与者随后将采取最佳策略。但是存在这样一种可能，博弈到达当前决策点的唯一途径是通过并不完美的博弈。我们需要用两种非常矛盾的想法进行这个实验：只有并不完美的博弈才能把这个博弈带到当前的决策点，但是我们假设以后的博弈是完美的。美国小说家斯科特·菲兹杰拉德（F.Scott Fitzgerald）曾经说过："对最高智力的测试是这样的一种能力，大脑中同时存在两种截然不同的观点却还能保持正常运转。"牢记这一思想，我们注意到，尽管逆向归纳法存在缺陷，就参与者该如何进行博弈及对当前参与者如何实际进行一场博弈而言，它还是一个很有说服力的原理。

练习题

1. 我们回过头来讨论第 2 章所描述的情形，伽利略有可能遭遇调查。让我们来记录下来究竟发生了什么。首先，罗马教皇乌尔班八世将伽利略告上法庭，伽利略于 1633 年 4 月 12 日接受审判。之后跟大多数审判一样，他遭到了来自委员们的口头劝说，伽利略承认在他所著的一本书中他的观点偏离了哥白尼学说（即使他并没有那样）。接着伽利略遭到了"目的调查"，包括提交刑讯文件。最终审判于 1633 年 6 月 22 日进行，69 岁的伽利略因身体状况欠佳请求法庭的宽恕。审查官以刑讯逼供及锒铛入狱相威胁，强迫伽利略放弃，诅咒，憎恶他的信仰和他的工作。伽利略屈服，最终被判处终身监禁并要苦修宗教。鉴于他的年龄（或许是考虑到他的名望），此次判决以在家软禁的方式执行。法庭允许他回到佛罗伦萨附近的别墅，他将在那里过完生命的最后几年。现在我们把这一案例模式化并开始讨论这个简单的模型。图 2—3 中广泛形式的博弈在图 8—18 中得以再现。

a. 找到所有的纳什均衡（暗示：先找到策略博弈。）

b. 找到所有子博弈精炼纳什均衡。

c. 对于并不是子博弈精炼纳什均衡的纳什均衡，解释一下为什么。

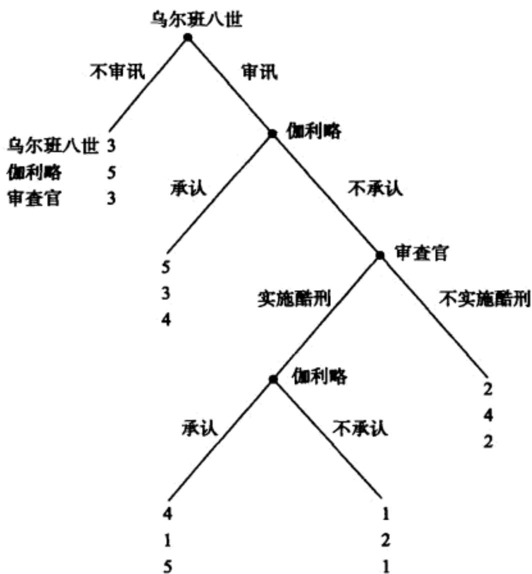


图 8—18 伽利略及审讯

2. 19 世纪 80 年代依然有侵权案件发生，只是大多发生在较隐蔽的地方而不是大庭广众之下。这些虚张声势的金融家忙于恶意收购，他们通过购买目标公司的股票，破坏目标公司既存的管理以得到这家公司（因此我们说这种收购是恶意的）。这种投资者就是我们所熟知的抢劫者，包括布恩，丹尼尔（T.Boone Pickens），詹姆士•歌德史密斯（Sir James Goldsmith），亨利•卡维斯（Henry Kravis），维克多•波斯纳（Victor Posner）。所有这些都是电影《华尔街》（*Wall Street*）中虚构的，其中道格拉斯扮演猎头戈登以支持"贪婪是有道德的"这一言论著称。这个时期到处是滑稽的行话，通过昂贵的财政组织结构，管理会消耗一种"毒药"，这将很难令这种恶意收购尽善尽美。在许多案例中，这种恶意的收购会遭到反抗，被收购者要求猎头公司收购其股票加付额外费用；这种额外收费就是平时所说的"绿票讹诈"，是敲诈勒索的收益。为了得到猎头和管理者之间策略的要点，我们要考虑图 8—19。猎头最初购买股票以找到哪种管理者决定了是否支付额外的费用购买股票（绿色敲诈费用）。如果不需要支付绿色敲诈费用，猎头决定是否购买更多的股票以便控制目标公司。

a. 找到子博弈精炼纳什均衡。

b. 找到一个不是子博弈精炼纳什均衡的纳什均衡，解释一下为什么它不是子博弈精炼纳什均衡。

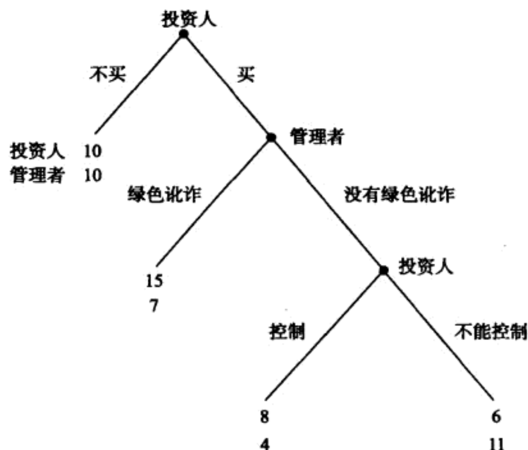


图 8—19 绿色敲诈

3. 我们回过头来讨论电影《赎金》（*Ransom*）中的绑架博弈（在第 2 章中首次被讨论），此博弈在图 8—20 中得以重现。求出所有子博弈精炼纳什均衡。

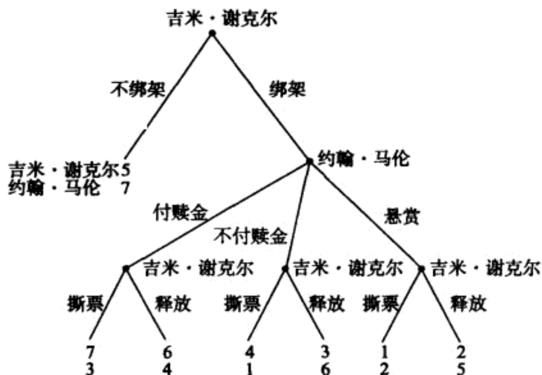


图 8—20 电影赎金的广泛形式

4. 本杰明•富兰克林曾说过："没有强制力的法律很难发挥作用，太严苛的法律很难得到执行。"为改变他的想法，图 8—21 中的博弈有三个参与者：立法者、公民、法官。这里有三种类型的法律：惩罚较轻、惩罚适中、惩罚严厉，立法者需要在这三种类型的法律中做出选择。作为对此次立法的回应，公民要决定是否遵守。如果他

有违法行为，法官决定是否宣判其有罪并予以惩罚。利用子博弈精炼纳什均衡，衡量支持了富兰克林的断言，导致立法者选择了惩罚适中的法律的不确定收益（用字母表示而不是用数字表示）的价值。

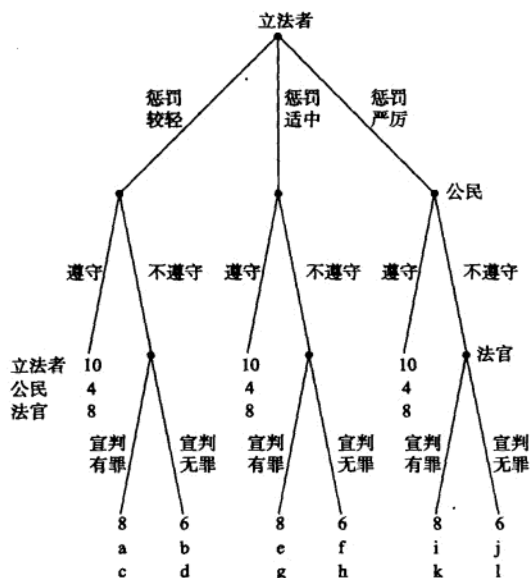


图 8—21　惩罚：从重抑或从轻

5. 诺贝尔获奖者托马斯·舍恩（Thomas Schelling）针对如何在盖伊的身份暴露后也能让其将被绑架者释放这一问题提出了解决方案。让我们回到绑架现场，用盖伊代表绑匪，用奥兰多代表被绑架者。问题在于一旦奥兰多看到盖伊的脸，奥兰多很有可能会被撕票灭口，因为如果奥兰多被释放，他将帮助警察通缉盖伊。这一情形正如图 8—22 所述。盖伊首先决定是否绑架奥兰多。奥兰多要决定是否透露一些自己的不为人知的犯罪细节给盖伊（可能是奥兰多挪用了教会的资金，或者对他的妻子隐瞒的一件大事。）接着盖伊决定是杀死奥兰多还是释放他。如果奥兰多被释放，他要决定是否告知警察盖伊的身份。如果奥兰多报警并且之前他已经将自己的秘密泄露给了盖伊，盖伊要决定是否与他人分享这个秘密。找到子博弈精炼纳什均衡，你将发现托马斯提议的解决方案。

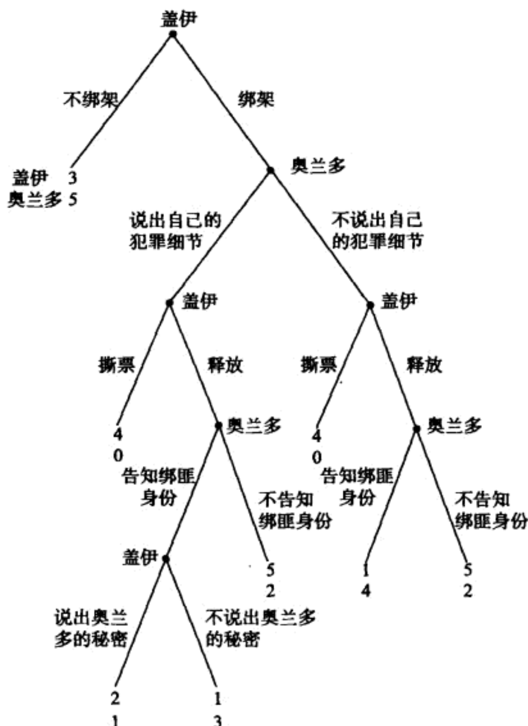


图 8—22　修改的绑架方案

6. 1842 年，伊利诺伊州首府斯普林菲尔德市的《Sangamo》杂志出版了一封批判伊利诺斯州审计员詹姆士·希尔兹（James Shields）的信件。尽管这封信的署名是丽贝卡，但希尔兹怀疑是州议员亚伯拉罕·林肯写的。正如图 8—23 所示，希尔兹考虑让林肯发起挑战，而且正如历史所记载的，他确实那样做了。作为回应，林肯要么逃避挑战，要么接受他的挑战，并且林肯有选择武器的权利。我们将允许林肯来决定是否提供道歉的种类。（请允许我在语法上做些颠倒。"道歉的种类"意思是经过讨论给希尔兹一个光荣撤退的机会——这是林肯的最终决定。）如果他决定迎接挑战，那么林肯面临四种选择：用枪做武器，用枪做武器且道歉，用剑做武器，用剑做武器且道歉。（谢尔兹是有名的神枪手，所以林肯选择骑兵用的最大号的砍刀作为武器，假设这个砍刀长 193 厘米，这样一来，对抗脾气较急躁的希尔兹，林肯占据了最大优势。）作为对这四种选择的回应，希尔兹必须决定是继续迎接挑战还是放弃。（在后面的案例中，如果实际上林肯道歉，谢尔兹接受他的道歉。）找到所

有子博弈精炼纳什均衡。作为最后的暗示，林肯曾说过："如果我与希尔兹之间存在矛盾时我的优异的表现能够被记住，那么我将不会被遗忘。"

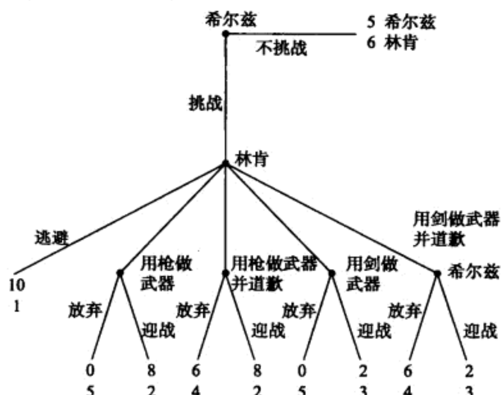


图 8—23　林肯与希尔兹之间的博弈

7. 尼克松总统第二次任职期间发生了一件丑闻，被称为"周六之夜大屠杀"。尽管没有人被解雇，但是许多人在联邦政府的地位确实受到了影响。尼克松政府试图让助手接近总统以掩饰罪行。作为侦查的一部分，司法部长艾略特·理查德森（Elliot Richardson, 不属于被隐藏的部分）把哈佛法学院的教授阿奇博尔德·考克斯（Archibald Cox）当成一个特殊的公诉人。尼克松总统非常担心考克斯的调查，他经过深思后命令理查德森解雇考克斯（在图 8—24 中表现为最初的决策点）。当尼克松把他的意图传达给理查德森时，后者传递出这样的意思，如果他解雇了考克斯，他可能会面对辞职的压力，但是或许他更愿意不解雇考克斯，这样他可能还是要辞职。关于解雇或者不解雇考克斯，辞职或者不辞职，理查德森存在四种可能的排列组合，扩展式博弈中描述了这四种排列组合。如果理查德森确实选择了辞职但不解雇考克斯，那么尼克松将面临摆脱考克斯的难题。如果理查德森选择既不解雇考克斯也不辞职，那么尼克松必须考虑是否解雇理查德森。一旦理查德森离开，首席检察官威廉·拉克尔肖斯（William Ruckelshaus）将军将成为代理人，他也跟理查德森一样将面临四种选择。如果拉克尔肖斯同样选择辞职但不解雇考克斯，那么罗伯特·勃克（Robert Bork）将变成代理人，他将面临四种同样的选择。为了简化问题，我们不考虑勃克，尽管理查德森和拉克尔肖斯一样都拒绝解雇考克斯并辞职后，勃克出场了，他解雇了考克斯却没有辞职。找到所有子博弈精炼纳什均衡。

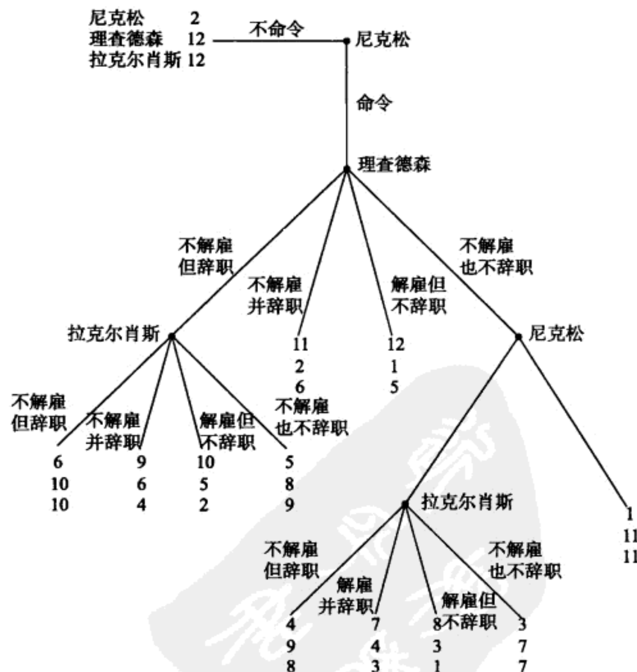


图 8—24

8. 七个小妖精正在决定如何撕碎 100 艘大型帆船。这几个妖精分别被记 A，B，C，D，E，F，G，他们被赋予了超能力，A 是最柔弱的而 G 是最强壮的。博弈从 A 开始，他对这 100 艘帆船提出了自己的分配计划，这个分配就是每个成员的所得是从 0 到 100 中间的数，其总和为 100。所有的成员同时投票，是赞成这个提议还是反对。如果至少一半的人投赞成票，那么博弈结束，这个提议通过。如果至少一半的成员反对，其他妖精会对 A 下咒语把他变成一个房子一周。此时就该 B 为剩余的六个成员制定分配计划了。同样的，如果至少半数的人同意，那么提议通过；如果未通过，B 被变成房子一周。接着是 C 上场。这个过程不断持续，直到任何一个提议被剩余的成员以半数或以上的票数通过或者除 G 之外的所有成员都被变成了房子，此时 G 得到了这 100 艘帆船。假设一个成员被变成房子后所得收益为-1 000，如果没有变成房子所得收益就是帆船的数量。运用子博弈精炼纳什均衡的概念，最后会发生什么呢？（重点关注这个子博弈精炼纳什均衡，如果一个参与者在赞成和反对之间持中立态度，那么他将对这个提议投反对票。）

9. 考虑一下图 8—25 所示的四个参与者的博弈。找出所有子博弈精炼纳什均衡。

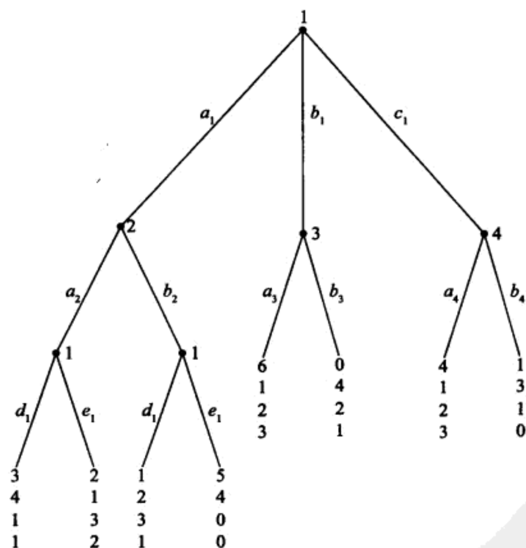


图 8—25

10. 托德和史蒂文的非常富有的伯父给他们留下了 100 公斤的黄金。他们分配这些财富的协议跟他们伯父的意志也有一定关系。达成一个协议共经历了三个回合。在奇数（偶数）回合，需要托德（史蒂文）提出一个提议。（托德在奇数回合提出建议，史蒂文在偶数回合提出建议难道不是明智的吗？）作为对这个提议的回应，另一个侄子可以接受也可以拒绝。如果他接受，那么博弈结束，提议通过。如果他拒绝，博弈进入下一回合。如果到第三个回合结束都尚未达成一致，那么所有的这些遗产将捐给慈善机构，托德和史蒂文都得不到这笔钱。此外，如果每一回合结束时协议未达成，那么全部遗产中的一小部分（$1-d$）将被捐给慈善机构，其中 $0<d<1$。所以在第二轮开始时有 $100d$ 公斤的黄金剩余（第一回合尚未达成协议）并且在第三回合开始时仅有 $100d^2$ 公斤的黄金（前两个回合都未达成协议）。换句话说，拖延达成协议的时间是要付出一定代价的，当然这种成本只有协议得不到通过时存在。每个侄子的收益就是他最终所得黄金的公斤数，所以不必在乎另一个侄子的收益，也不用关心他们伯父所热衷的慈善事业。为了对提议做某种标记，假设第 t 个回合根据提议所得的收益为 x_t，其中 x_t 是托德所占有的剩余的那部分黄金，从而史蒂文所占有的部分是 $1-x_t$，$0 \leq x_t \leq 1$，所以 x_t 包括从 0 到 1 的任何数据。找出一个子博弈精炼纳什均衡。

11. 思考一下以下出自《午夜善恶花园》（*Midnight in the Garden of Good and Evil*）的内容：

> 有一位女士，她是上层社会的贵妇，东南部最富有的富豪之一，叫做萨万娜。她有一个铜矿。她在城市的豪华住宅区建造了一处豪宅，这座豪宅模仿了著名的路易斯安那农场的设计，有巨型白色圆柱和曲线型楼梯。从水中可以看到倒影。每个路过的行人都会驻足欣赏。我很羡慕她。对我而言，她曾经就像一位母亲一样。但是这个豪宅的成本非常低。许多年前，她为自己的住宅定制了两扇铁门。这两扇门是专门为她设计制作的。但是当这个作品出现在她面前时，她表现出异常的厌恶，她说这是令人生厌的、肮脏的东西。她喊着“把它们拿走，我再也不想看到。”接着她撕毁了账单，1 400 美元在那个时代是一个不小的数目。制造商把门带了回去，但是他们并不知道该如何处理，

毕竟没有多少人需要这种装饰用的巨大的铁门。他们唯一能做的就是当做废铁卖掉，所以他们把价格从 1 400 美元降到了 190 美元。很自然的，接下来这位女士派出一个人拿着 190 美元从制造商那里买到了这两扇门，今天它们就安在那位女士豪宅的门柱上，而这也是它们被设计的原始地。

使用逆向归纳法，你能解释制造商错在哪里了吗？

12. 图 8—26 重新介绍了第 2 章所述的讨价还价博弈。分析所有子博弈精炼纳什均衡以便找出一个参与者任何时候都能接受的最佳行动。也就是说，如果选择接受或者选择其他行动能够使参与者获得最大收益，那么参与者会选择接受。

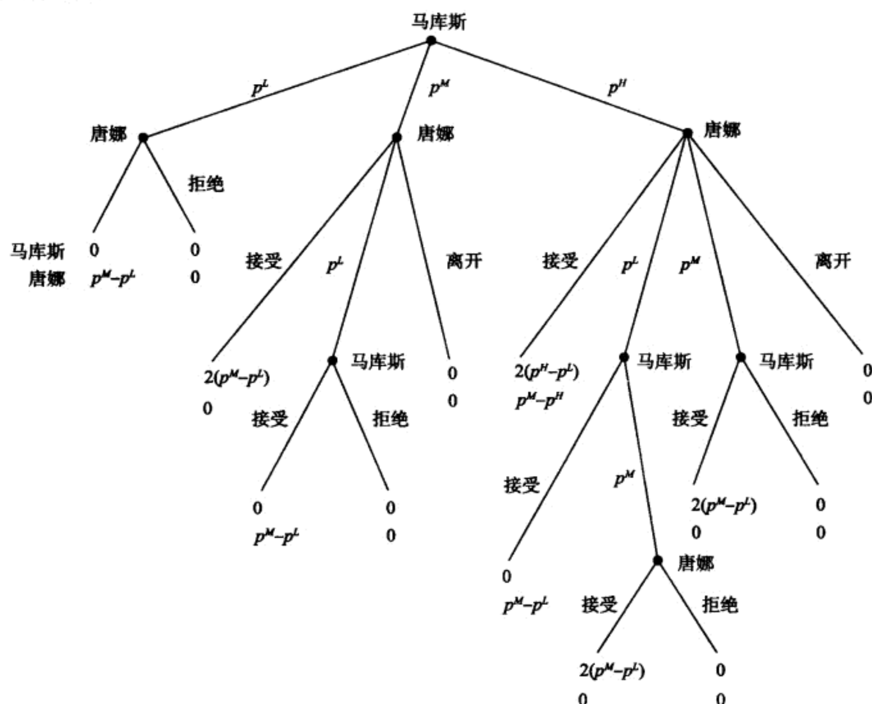


图 8—26　讨价还价博弈

第9章

黑暗中的交替：不完美信息下的序列博弈

学习目标

9.1 子博弈精炼纳什均衡
9.2 利用相关均衡收益代替子博弈
9.3 承诺

本章导读

现在我们来关注一下不完美信息博弈。如前文所述，**当参与者需要做出决策的时候往往并不完全清楚已发生的博弈历史，这一类型的博弈我们称之为不完美信息博弈**。用专业术语来说，在一场不完美信息博弈当中，许多信息集包含不止一个决策节点。与我们在第8章了解到的相似，由于其中有相信不可置信的威胁的参与者（即参与者发出的威胁违背了自己的利益最大化），纳什均衡对于不完美信息博弈来说是不切实际的。对于完美信息博弈而言，我们要找出一个有说服力的求解方案，其目的在于剔除掉包含于纳什均衡当中的不可置信的威胁。

方法正如第8章所提到的：扩展式策略能够适用于不完美信息博弈。由于不完美信息博弈当中信息的构成更加丰富，导出解的过程也有些复杂。我们首先用示意图来阐明这种方法，在9.1小节当中再对这一方法进行更加正式的概括。对于需要复习一下不完美信息博弈的读者，在往下进行之前可以回顾一下2.2小节中的内容。

让我们来回顾一个绑架博弈。绑架者盖伊决定是否释放受害人奥兰多，与受害人的亲属薇薇卡决定是否付赎金，该二者的行为几乎是同时发生的。图9—1重现了第2章中的扩展式博弈。

如果把在第8章当中介绍过的公式用在这里，我们就需要从绑架者的最后两个决策节点中推导出最优行为。但是，仔细考虑以下难点：与盖伊的绑架行为和薇薇卡的付赎金的行为相互关联的节点Ⅲ，与节点Ⅳ都是信息集的一部分。因此盖伊并不清楚自己正处于节点Ⅲ还是节点Ⅳ（也就是说，他不知道薇薇卡有没有付赎金），究竟什么才是理性的信念，以及自己在博弈当中处于什么位置都无法知晓。

仔细观察图9—2，为了能更好地理解所面临的难点，我们把正在研究的博弈过程截取了一部分。这段截取下来的博弈过程开始于非单元素集合的信息集。此博弈开始时透露了些信息（薇薇卡的行动），但我们不清楚刚才发生了什么。博弈论不能很好地解释这类含糊的问题，而图9—2有别于我们以往所了解的任何博弈类型。

PDG

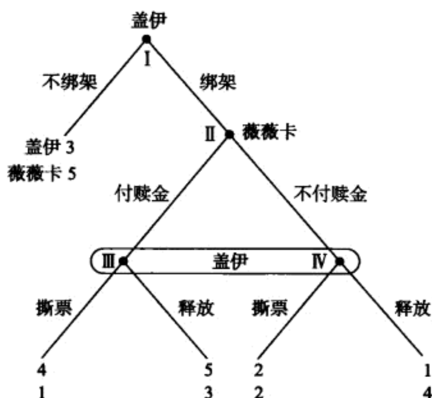


图 9—1　在绑架博弈中，付赎金和释放受害人的决定几乎是同时做出的

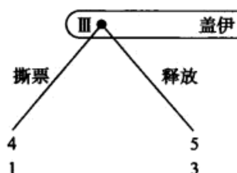


图 9—2　盖伊在节点 III 时的局势

现阶段，我们只好放弃努力，同时承认无法解出始于盖伊的信息集的行为。承认失败，我们沿着博弈树爬到节点 II 处薇薇卡的信息集。由于节点 II 是一个单元素集合，刚才讨论的问题都没有发生。同时，由于没有解出盖伊随后的行为（他究竟是杀害还是释放了受害人），要知道，求出薇薇卡的行为需要同时求出盖伊的行为。这并不会造成问题，因为这一部分的博弈过程，如图 9—3，是可以辨别的。

图 9—3 中博弈的策略形式用于帮助我们求出图 9—4 当中的纳什均衡。这个博弈包含了一个特殊的纳什均衡：薇薇卡不付赎金，盖伊杀死奥兰多。因此，假如存在一个绑架，如果赎金没交，奥兰多就会被杀。

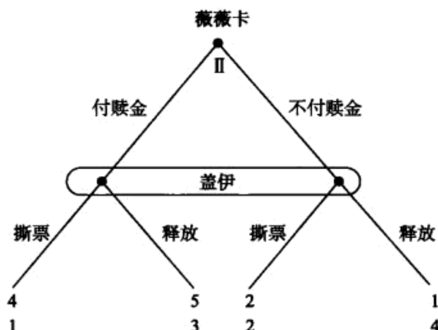


图 9—3　薇薇卡和实施绑架后的盖伊所面对的局势

薇薇卡

		付赎金	不付赎金
盖伊	撕票	4,1	2,2
	释放	5,3	1,4

图 9—4　薇薇卡和盖伊面对的策略形式

让我们回到第一个节点（此时盖伊正在策划一场绑架），我们看到如果盖伊没有收到赎金他就会杀掉奥兰多。如果他实施了绑架他的期望收益是 2，图 9—5 描述了他所面对的形势。很明显，他不会选择绑架奥兰多，于是得到的收益是 3。通过之前的分析，图 9—1 中博弈的解是一个策略组合（不实施绑架 / 撕票，不付赎金）。



图 9—5　当盖伊求出实施绑架后的结果以后所面对的形势

PDG

9.1 子博弈精炼纳什均衡

在这里我们提供一种比本章导读中使用的方法更为正式和概括的描述。这种方法的核心是辨别出博弈中那些“看上去像博弈”的部分，对于这些博弈，我们可以分别解出并要求参与者对每个“博弈中的博弈”都按照最优选择行动（在某种意义上他们的博弈符合纳什均衡）。

回顾前一章。对于完美信息博弈，子博弈精炼纳什均衡要求一个参与者对于每个信息集都按照最优选择行动。在按最优选择行动的博弈当中，假设其他参与者也采取同样的行动。如图 9—1 所示，当一个博弈是不完美信息博弈时，这个步骤不能被用作任何信息集的开始。我们不能从一个非单元素信息集开始。那么我们是否能这样总结——我们可以分析出博弈的任何开始于单元素信息集的部分？遗憾的是，即使单元素信息集也会造成难题，如图 9—6 中博弈的情况。

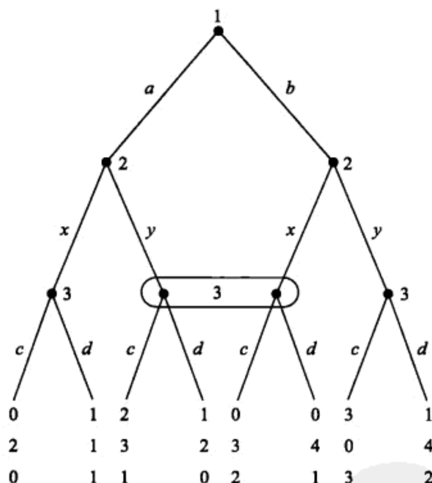


图 9—6　扩展式博弈

注意，参与者 1 选择从 a 行动开始时的情况。那样的话参与者 2 必须要在 x 和 y 之间做选择，同时参与者 3 必须在 c 和 d 之间选择。问题在于，如果参与者 2 选择了 y，那么参与者 3 该如何预测自己所处博弈的位置呢？她无法辨别是通向 y 的 a 路径，还是通向 x 的 b 路径。图 9—7 截取了此博弈的一部分使得这个意义不明确的博弈更加清楚。我们也不知道参与者 3 应该如何选择行动，因为参与者 3 的信念以及对于自己在博弈中所处的位置并不明确。由于参与者 3 的行为选择不明确，那么参与者 2 应该如何选择就很难说了。上述所定义的某个扩展式博弈的一部分本身就是一个意义明确的博弈。考虑到扩展式博弈的不同构成，这需要用到一些形式体系。

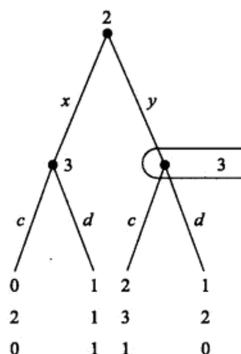


图 9—7　扩展式博弈中含义不明确的部分

我们把一个子树以及它所包含的所有节点称为非终端节点。图 9—1 中的绑架博弈有 4 个子树，分别在图 9—8 中被圈出来。**一个标准子树特指包含所有子树当中存在至少一个节点的全部信息集。**包含一个信息集的一个节点的标准子树，必须包含这个信息集中的所有节点。图 9—1 中的博弈有两个标准子树，它们在图 9—9 中被圈了出来。图 9—8 中的子树 3 不是一个标准子树，因为它包含了节点Ⅲ但没包含节点Ⅳ，也就是说它包含了一部分，但没有包含盖伊的信息集中的全部节点。同样，子树 4 也不是一个标准子树。子博弈和其对应的一组收益就是一个标准子树。

子博弈与原博弈看上去很相似，它是“博弈中的博弈”。避免不明确博弈的关键之处（比如图 9—2 和图 9—7），在于集中关注博弈当中被层层包围的部分。如果它包含了某个信息集的一个节点，那么也包含了这个信息集的所有节点。要找到一个子博弈，我们只需要从一个最终决策节点开始，沿着博弈树向上直到符合要求的前述条件出现。

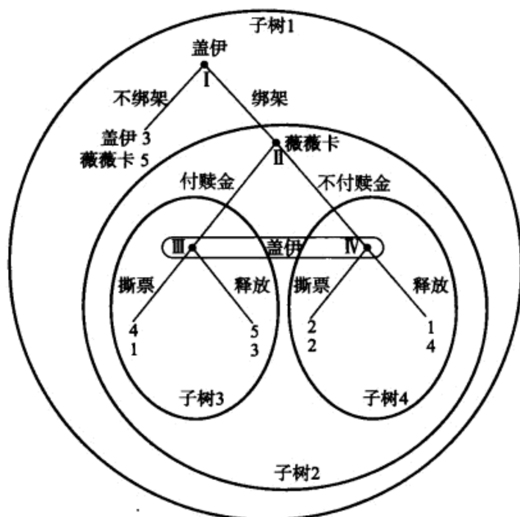


图 9—8　绑架博弈的四个子树

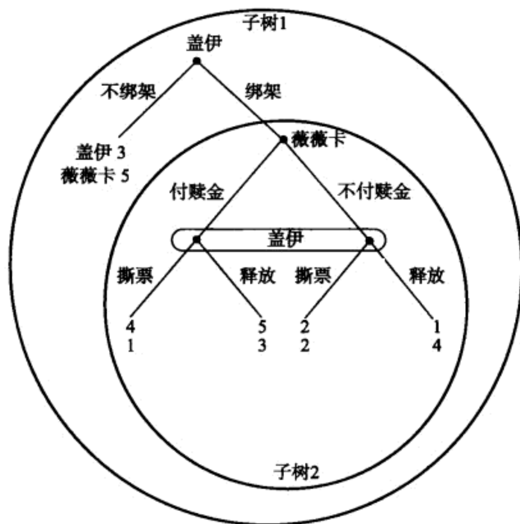


图 9—9　绑架博弈的两个标准子树

9.1 理解测验

图9—6所示的博弈中有几个子树，又有几个标准子树？

一个符合子博弈精炼纳什均衡的策略组合，必须使得纳什均衡适用于每个子树。为了更加精确，我们把一个子博弈的子策略定义为一组策略中的某一部分，这部分只规定了这个子博弈的信息集中的行为。仔细思考图9—10中突出显示部分的策略组合（绑架／释放，付赎金），这个策略组合中的行动引出了此博弈的三个信息集。

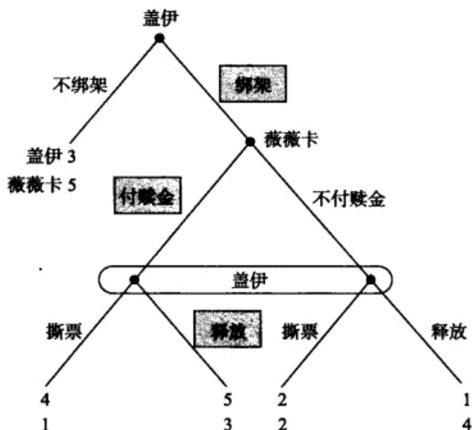


图 9—10　策略组合（绑架／释放，付赎金）中的行动引出博弈的三个信息集

这个策略组合的子策略属于从薇薇卡的信息集开始的子博弈（如图9—3所示），只包括那些与子博弈相关的行动。这样，子策略组合就是（释放，付赎金），盖伊在他的信息集中选择释放，薇薇卡在她的信息集中选择付赎金。

> **子博弈精炼纳什均衡（Subgame Perfect Nash Eguilibrium）：如果策略组合的每个子博弈的子策略组合都是一个纳什均衡，那么这个这个策略组合就是一个子博弈精炼纳什均衡。**

这个概念非常简单。一个策略组合蕴含着每个子博弈特定的应变计划，每个子博弈都是“博弈中的博弈”。对于每一个这样的子博弈而言，我们可以提出这样的问题：在所有参与者都按照他们的最大化利益选择行动的前提下，包含的应变计划是否合理，也就是说，这些计划能否构成一个纳什均衡。一个子博弈精炼纳什均衡要求每一个子博弈均构成均衡。注意，这一定义适用于所有的扩展性博弈，包括完美信息博弈和不完美信息博弈。

逆向归纳法从解决最终子博弈的纳什均衡行为开始着手，然后用纳什均衡的收益代替每个子博弈，借此定义一个新的（同时也是更小的）扩展式博弈。在这个新的博弈中，你将实施相同的两个步骤：对最终子博弈求出纳什均衡解，然后用纳什均衡的收益来代替每个子博弈。继续沿着博弈树向上追溯，直到到达初始节点。一个接一个的子博弈以此方式构成了子博弈精炼纳什均衡。

案例研究

英国情报机关

为使这些抽象的概念具体化，我们来浏览一个例子。如果你曾经读过任何一本有关邦德（James Bond）的小说，那你应该知道，英国情报机关的领导是一位叫做“M”的神秘人物，而一个类别为“00”的特工需要奉命执行“杀人任务”。当 M 必须选择一位特工去执行一件极其关键而又危险的任务时，我们来仔细思考一下她所面对的情况。如图 9—11 所示，她可以将任务指派给特工 003，也可以指派给特工 007，或者同时交给他们两个。不管是谁接到任务，都必须在执行任务时，在“专注”与“热忱”之间做出选择，而其中后者意味着愿意为之牺牲生命。

为了求出这个博弈的解，需要先从当只有特工 007 被 M 选中时，他所面对的信息集开始着手。首先应该明确的是，热忱是特工 007 执行任务时的最佳选择，所以我们把这个子博弈用纳什均衡中的收益（5，4，3）来代替。

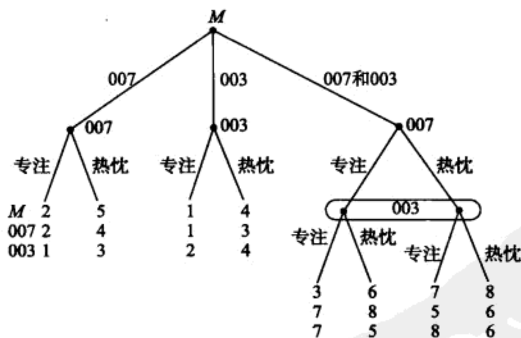


图 9—11　英国情报机关博弈

接下来考虑当只有特工 003 被指派执行任务时的子博弈。同样的，特工 003 也选择了热忱，因此这个子博弈可以用收益（4，3，4）来代替。

第三个子博弈的推导是当 M 把任务同时指派给两位特工时的情况。特工 003 和特工 007 所面对的此博弈的策略形势如图 9—12 所示（在这里我们不管 M 的收益，因为它们与剔除掉此博弈中的行为选择无关）。这个博弈中纳什均衡的独特之处在于两位特工都选择了专注。收益背后的逻辑是这样的：每个特工在任务仅仅被指派给他一人的情况下都愿意选择热忱（正如我们在分析前两个子博弈的时候所发现的），但如果两位特工在被同时指派任务的情况下，每个人都认为选择专注去执行任务足矣。这个子博弈可以用收益（3，7，7）来代替。

		特工 003	
		专注	热忱
特工 007	专注	7,7	8,5
	热忱	5,8	6,6

图 9—12　当 M 选择把任务同时交给特工 003 和特工 007 时，二人所面对的策略式博弈

我们已经求出三个子博弈的解，并用与之相关联的纳什均衡收益值代替这些解，那么 M 现在所面对的情况可以用图 9—13 来描述。对于这个只有一个参与者的博弈，纳什均衡是只把任务指派给特工 007。从 M 的收益值中可以体现出来，相较于特工 003 她更倾向于选择特工 007（可能后者更加熟练），而且相对于让两位特工以较弱的水准执行任务，她更希望只让一位特工执行任务足够热忱。



图 9—13　在三个子博弈已经通过均衡选项解出时，M 所面临的情况

总而言之，这个独特的子博弈精炼纳什均衡可以用下面的策略组合描述如下：M 指派特工 007 去执行任务，如果任务仅仅指派给他一人，他执行任务时选择热忱；

如果任务被同时指派给特工003，那么他选择专注。而如果任务仅仅指派给特工003，他执行任务时选择热忱，如果任务被同时指派给两个人，那么他选择专注。

如果我们能同时说明，一个纳什均衡什么时候不能构成一个子博弈精炼纳什均衡，也许这会对理解子博弈精炼纳什均衡的概念有所帮助。为此，考虑策略组合（007和003，专注 / 专注，专注 / 专注），这是一个纳什均衡，但不是子博弈精炼纳什均衡。让我们先来看看它为什么是一个纳什均衡。

在这个策略组合当中，M的收益是3，如果她只选择特工007那么她的收益就会是2。因为无论007是不是唯一被选中的特工，他都会选择专注。如果她只选择特工003那么她的收益将会是1，因为003会选择专注。这样的话，M同时选择两个特工的确是最优选择。现在，如果M同时选择两人，那么007选择专注 / 专注的收益是7，选择热忱 / 专注的收益是7（因为当M同时指派两人，它要求同样的行动），不管是专注 / 热忱还是热忱 / 热忱的收益都是5。这样，特工007的策略也是最优的。与之相似，我们也能说明特工003相对于专注 / 热忱和热忱 / 热忱，严格倾向于选择专注 / 专注，尽管热忱 / 专注产出同样的收益。综上所述，因为每个参与者的策略都是相对于其他两个参与者选定策略的最优选择，所以（007和003，专注 / 专注，专注 / 专注）是一个纳什均衡。

为了说明（007和003，专注 / 专注，专注 / 专注）不是子博弈精炼纳什均衡，我们只需要找到一个子博弈，这个子博弈的子策略组合（007和003，专注 / 专注，专注 / 专注）不是纳什均衡。为此，考虑当M只把任务指派给特工003时的子博弈。这是一个只包括特工003一个参与者的博弈，与之相关的子策略专注 / 专注迫使他选择专注。然而，专注产出的收益为2，而选择热忱带来的收益是4。换句话来说，在被指派为唯一的特工的事件中，特工003的策略使得他选择专注。但事实上，热忱才是他的最优选择。

最后要注意的是，只有当M相信如果仅指派一个特工，他们不会选择热忱时，同时指派两个特工才是他的最佳选择。但如果M相信仅指派特工007一个人这项工作，选择热忱是他的最佳选择时，M会将工作指派给特工007一个人。事实上，这就是我们推导出的独特的子博弈精炼纳什均衡。

9.2 理解测验

回顾第2章中的抢劫博弈（如图9—14所示）。抢匪先是选择是否携枪，如果带了枪，是否向受害人西蒙出示。无论是否看到了枪，西蒙的选择是反抗或者放弃。找出所有的子博弈精炼纳什均衡。

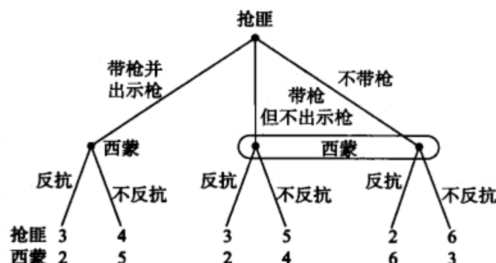


图9—14 抢劫博弈

9.2 利用相关均衡收益代替子博弈

在分析下列几个例子之前，我们首先简单思考一下，不完美信息博弈中是否总存在一个子博弈精炼纳什均衡，如果是，是否存在多个这样的均衡。回顾第8章，我们知道每个完美信息博弈中都存在子博弈精炼纳什均衡。遗憾的是，对于不完美信息博弈来说，我们没有得出这样的结论。如果一个子博弈不存在任何纳什均衡（在纯策略中），就一定不存在子博弈精炼纳什均衡（在纯策略中）。然而，与第7章中提到的内容相似，如果我们允许参与者随机选择，那么在混合策略中存在一个子博弈精炼纳什均衡。

在有关均衡数量的问题上，我们知道当在某些决策节点上两种行动产生相同的最大收益时，一个完美信息博弈中存在多个子博弈精炼纳什均衡。在不完美信息博弈中，均衡多样性的问题更突出。即使当所有行动产生的收益都不同（如我们将要分析的OS/2博弈），也可能存在多个子博弈精炼纳什均衡。因为在第4章和第5章中，我们已经为许多博弈求得多个纳什均衡，所以，对我们来说，这类情况很普遍。

实际上，如果存在多个子博弈精炼纳什均衡，重复利用上一小节提到的方法，我们可以推导出这些均衡。所以，**如果一个子博弈满足存在多个纳什均衡的条件，**

我们就应该选择其中一个均衡，用相关均衡收益来代替子博弈，然后利用这个方法一直研究到树的首节点。这个过程只验证了子博弈精炼纳什均衡中的其中一个均衡，你还需要选择另外一个均衡，重复相同的过程。下面的这个例子说明了这个过程是如何完成的。

◯案例研究

OS/2

1981 年，当 IBM 决定要进军计算机市场时，公司就决定要使 IBM 的产品成为行业的标准。因为当时市场上已有家用电脑销售，IBM 决定在其他家用电脑成为行业标准之前，快速进入电脑市场。IBM 并非只开发个人电脑，它还外包其他关键领域的开发，如微处理器和操作系统。在操作系统的开发上，IBM 寻求与一家小公司合作，而这家小公司正是由保罗•艾伦（Paul Allen）与比尔•盖茨共同运营的微软。

为了满足 IBM 的需求，微软以 50 000 美元从西雅图电脑公司购买了 QDOS（“Quick and Dirty Operating System”），在此操作系统上稍加修改，把终端产品 MS—DOS 卖给了 IBM 公司。IBM 允许微软保留 MS—DOS 的专利权，这被认为是 20 世纪最糟糕的商业决策。当时 IBM 预测是公司从个人电脑的销售中获利，而非从操作系统中盈利。然而，历史证明情况并非如此：个人电脑仅仅是一件产品，一件许多竞争者很容易就生产出高度相似产品的产品；而微软凭借操作系统在网络产生的效益在软件市场上独领风骚。

一旦意识到这个错误，IBM 开始开发自己的操作系统——首先与微软联合，1987 年发布了 OS/2 操作系统。那时，微软已发布了 Windows 的早期版本，主要竞争对手是 OS/2 系统。尽管 IBM 是个人电脑的主要生产者，尽管 OS/2 比 Windows 运行速度更快更稳定，大多数的消费者还是不愿意购买 OS/2 系统的个人电脑，原因在于只有少数软件才能在此系统上运行，而那时 Windows 操作系统支持大部分软件的运行。当 Windows 操作系统已经兼容大部分软件时，对 IBM 公司来说，推销 OS/2 就是一个巨大的挑战。

为了分析 IBM 是否要开发自己的操作系统这个战略场景，思考图 9—15 中的扩展式博弈。IBM 首先决定是否要开发 OS/2。如果它选择开发 OS/2，那么假设 3 个公司同时决定是否开发 OS/2 系统支持的软件，隐含在收益中的情况表明，如果至少有两个公司开发此类软件，OS/2 将会有很大的市场。所以，OS/2 开发是否成功取决于是否有足够的兼容软件被开发。

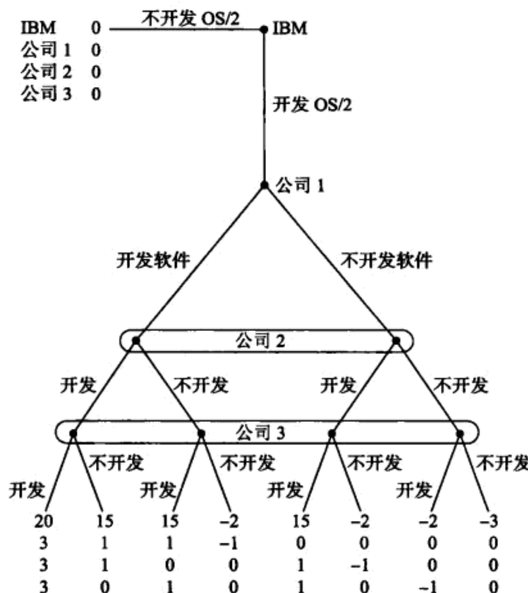


图 9—15 OS/2 博弈

为了求解这个博弈的子博弈精炼纳什均衡，首先思考当 IBM 选择开发 OS/2 时，3 个参与者子博弈的情况。博弈的战略表达式如图 9—16 所示。每个单元格的第一个数字是软件公司 1 得到的收益，第二个数字是软件公司 2 得到的收益，最后一个数字是软件公司 3 得到的收益。因为 IBM 公司在这个子博弈中不做出任何决定，所以略去 IBM 公司得到的收益。

公司 3，开发

		公司 2	
		开发	不开发
公司 1	开发	3,3,3	1,0,1
	不开发	0,1,1	0,0,−1

公司 3，不开发

		公司 2	
		开发	不开发
公司 1	开发	1,1,0	−1,0,0
	不开发	0,−1,0	0,0,0

图 9—16 IBM 开发 OS/2 后，软件公司 1、2、3 面对的情况

PDG

存在两个纳什均衡（在纯策略中），一个是3个公司都开发OS/2系统下兼容的软件，另一个是都不开发此类软件。

在这个子博弈中，存在2个纳什均衡，要推导所有的子博弈精炼纳什均衡就需要求解这个博弈两次。首先假设在纳什均衡中，当3个公司都开发OS/2系统下兼容的软件时，IBM面临的情况如图9—17所示。它可以选择不开发OS/2，这样得到的收益为0，或者选择开发，得到的收益为20。

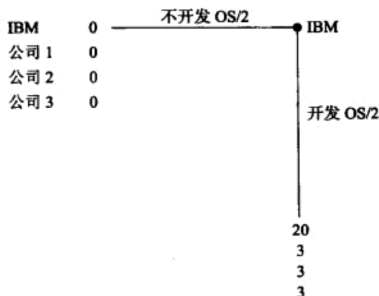


图9—17　当3个公司都开发OS/2系统下兼容的软件时，IBM公司面临的情况

很明显，IBM的最佳决策是开发OS/2系统，于是策略组合（开发OS/2，开发，开发，开发）是一个子博弈精炼纳什均衡。

现在思考另一个均衡，在这个均衡中，3个公司都不开发OS/2系统下兼容的软件，IBM公司面临的情况如图9—18所示。这个均衡使得IBM做出不开发OS/2的决定。这就为我们提供了第二个子博弈精炼纳什均衡：（不开发OS/2，不开发，不开发，不开发）。

两个子博弈精炼纳什均衡可以产生两个截然不同的结果：一个是IBM开发OS/2，且期望3个公司都开发兼容的软件；另一个是IBM放弃这个操作系统项目，因为它意料到3个软件公司都不开发相兼容的软件。

实际情况如何呢？IBM引入OS/2系统，但历史证明OS/2能兼容的软件很少。尽管IBM公司在电脑市场上占据很大份额，但也不能撼动微软公司Windows系统的主导地位。

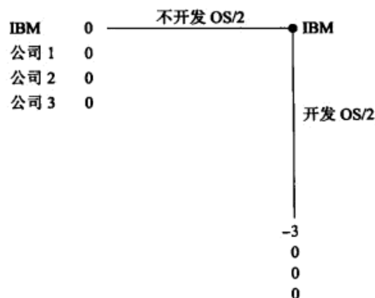


图9—18　当3个公司都不开发OS/2系统下兼容的软件时，IBM公司面临的情况

假设在这3个软件开发公司的子博弈中，存在一个对称纳什均衡，在均衡中，每个公司随机选择。令d表示公司选择开发软件的概率。对软件公司来说，要求的最佳随机选择方案得到的收益必须与它采取纯策略——开发或不开发得到的收益相等。假设公司2和公司3选择开发软件的概率为d，于是，这两个公司都开发软件的概率为d^2，这两个公司都不开发软件的概率为$(1-d)^2$；一个公司开发，另一个不开发的概率为$2d(1-d)$。[①]那么，公司1从开发软件中得到的收益为：

$$d^2\times3+2d(1-d)\times1+(1-d)^2\times(-1)=4d-1$$

把这个不等式和0（不开发软件得到的收益）划等号，我们得到$d=\frac{1}{4}$，所以，这3个软件公司中的任何一个公司选择开发的概率为$\frac{1}{4}$时，为一个对称纳什均衡。

给定这个子博弈的纳什均衡，IBM公司从开发OS/2系统得到的期望收益为：[②]

$$\left(\frac{1}{4}\right)^3\times20+3\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)\times15+3\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^2\times(-2)+\left(\frac{3}{4}\right)^3\times(-3)=\frac{20}{64}$$

因为不开发OS/2得到的收益为0，IBM发现开发OS/2系统是最佳方案。然而在开发过程中，IBM不能确定软件公司是否会开发相兼容的软件。如果至少有两个公司选择开发与OS/2系统相兼容的软件（事件发生的概率为$\frac{10}{64}$），那么OS/2系统就会占有市场份额。[③]但少

① 公司2开发软件，公司3不开软件的概率为d（1–d），而公司2不开发，公司3开发的概率为d（1–d）。这样，2家公司中有1家开发的概率为之间两概率的和，即2d（1–d）。

② 两家公司开发软件的概率是以下概率之和，公司1和公司2开发的概率，公司1和公司3开发的概率以及公司2和公司3开发的概率。其中每个概率均为$(1/4)^2(3/4)$，概率之和为$3(1/4)^2(3/4)$。

③ 3家公司写申请的概率为（1/4）3=1/64，2家公司写申请的概率为3（1/4）2（3/4），故概率之和为10/64。

于两个公司开发相兼容软件的概率为$\frac{50}{64}$，OS/2 系统被淘汰。正如我们所了解的，历史证明 OS/2 系统被淘汰了。

参议院的议程决定

美国参议院正考虑一次投票的 3 个方案，我们用 A、B 和 C 表示。100 个议员的偏好情况如表 9—1 所示。10 个议员倾向方案 *A*，同时把方案 *B* 作为他们的第二个选择。44 个议员把方案 *B* 作为他们的第一选择，同时把方案 *C* 作为第二选择。最后，46 个议员把方案 *C* 作为第一选择，同时把方案 *A* 作为第二选择。

表 9—1　议员的偏好

议员的类型	第一选择	第二选择	第三选择	议员的数量
Ⅰ	*A*	*B*	*C*	10
Ⅱ	*B*	*C*	*A*	44
Ⅲ	*C*	*A*	*B*	46

投票通过的步骤有三步：

步骤 1：参议院的领导人员通过选择两种方案设置议程。他的选择可以是 *A* 和 *B*、*A* 和 *C* 或 *B* 和 *C*；

步骤 2：给定两种方案，100 个议员同时投票；

步骤 3：步骤 2 中得票最多的方案和第三种方案结合，议员对这两种方案再投票，得票最多的方案通过。

例如，假定参议院领导人员选择方案 *A* 和 *C*，那么在步骤 2 中，每个议员投票 *A* 或 *C*。如果方案 *A* 得票更多，那么在步骤 3 中，议员对方案 *A* 和方案 *B* 再投票。如果在步骤 3 中方案 *A* 得到大多数议员的支持，那么方案 *A* 通过，而如果方案 *B* 得到大多数议员的支持，那么方案 *B* 成为法律的一部分。因为有 100 个参与者，树型结构很难构建扩展式博弈，所以我们还是用语言来描述这个过程。最后，假如议员最偏爱的方案被通过，他得到的收益为 2；如果他的第二选择被通过，得到收益为 1；如果他的第三方案被通过，得到的收益为 0。

我们关注的焦点是推导子博弈精炼纳什均衡，使议员真实投票成为最优策略；也就是说，在对两种方案投票时，议员会给他认为更好的方案投票。比如，如果进行方案 *A* 和方案 *C* 的投票，那么Ⅰ类型议员投票 *A* 方案，Ⅱ和Ⅲ类型的议员投票 *C* 方案。

步骤 3 的子博弈使得 100 个议员同时对两种方案投票（一种方案是步骤 2 中获胜的方案，一种是步骤 1 中不包含的那种方案），这些最终的子博弈由步骤 1 和步骤 2 决定，这与两个因素相关：参议院领导人员首选的两种方案，以及议员对两种方案投票的结果。尽管存在许多类似的子博弈，在推导一个均衡时起关键作用的是了解步骤 1 提出的两种方案，和步骤 2 中得票更多的方案（为了简化过程，我们省略得票相同的情况），那么：

- 思考步骤 3 中的子博弈，其中参议院领导人员在步骤 1 中提议方案 *A* 和方案 *B*。

  如果步骤 2 中方案 *A* 得票更多。于是议员在步骤 3 中必须选择方案 *A* 或方案 *C*，通过真实投票，方案 *C* 获通过，因为 90 个议员都偏爱方案 *C*（参照表 9—1），注意真实投票是最佳策略，因为获通过的方案不是由单个议员的投票结果决定。

  如果步骤 2 中方案 *B* 得票更多。[1]于是议员必须选择方案 *B* 或方案 *C*，通过真实投票，方案 *B* 获通过，因为 54 个议员都偏爱方案 *B*。
- 思考步骤 3 中的子博弈，其中参议员领导人员在步骤 1 中提议方案 *A* 和方案 *C*。

  如果步骤 2 中方案 *A* 得票更多。于是，在步骤 3 中通过真实投票，方案 *A* 以 56 : 44 获通过。

  如果步骤 2 中方案 *C* 得票更多。于是，在步骤 3 中通过真实投票，方案 *B* 以 54 : 46 获通过。
- 思考步骤 3 中的子博弈，其中参议员领导人员在步骤 1 中提议方案 *B* 和方案 *C*。

  如果步骤 2 中方案 *B* 得票更多。于是，在步骤 3 中通过真实投票，方案 *A* 以 56 : 44 获通过。

  如果步骤 2 中方案 *C* 得票更多。于是，在步骤 3 中通过真实投票，方案 *C* 以 90 : 10 获通过。

① 在步骤2中参议员依据个人最优方案投票，那么方案*B*的得票不会高于方案*A*。但请记得，我们必须确保该博弈中每个子博弈符合纳什均衡。

利用步骤3中的子博弈的纳什均衡结果，我们返回分析步骤2的子博弈。前面已经分析过，真实投票是最佳策略。所以，如果参议院领导人员首选方案*A*和方案*B*，那么方案*A*以56票获胜；如果他首选方案*A*和方案*C*，那么方案*C*以90票获胜；如果他首选方案*B*和方案*C*，那么方案*B*以54票获胜。

再利用这些分析结果，我们现在分析步骤1中的博弈，也就是参议院领导人员如何选择两种方案，在接下来的分析过程中，结果取决于他的选择：

- 如果在步骤1中提议方案*A*和方案*B*，那么方案*A*以56:44获胜（已证明）。于是，在步骤3（或在第二轮投票中），方案*A*和方案*C*竞争，方案*C*以90:10获胜（已证明）。所以，如果参议院领导人员在议程上首先提议方案*A*和方案*B*，方案*C*最终获胜。
- 如果在步骤1中提议方案*A*和方案*C*，那么方案*C*以90:10获胜。于是，在第二轮投票中，方案*C*和方案*B*竞争，方案*B*以54:46获胜。所以，如果参议院领导人员在议程上首先提议方案*A*和方案*C*，方案*B*最终获胜。
- 如果在步骤1中提议方案*B*和方案*C*，那么方案*B*以54:46获胜。于是，在第二轮投票中，方案*B*和方案*A*竞争，方案*A*以56:44获胜。所以，如果参议院领导人员在议程上首先提议方案*B*和方案*C*，方案*A*最终获胜。

参议院领导人员的最佳行为取决于他自己的偏好，如果他和Ⅰ类型议员有同样的偏好，那么他最希望方案*A*获通过，这样他就会首选方案*B*和方案*C*，结果方案*C*落选，方案*B*在步骤3中与方案*A*的竞争中落选。相反，如果参议院领导人员和类型Ⅱ的议员有同样的偏好，那么他就希望在步骤2中方案*A*和方案*C*竞争。尽管方案*C*将会获胜，但会在第二轮投票输给方案*B*；而如果参议院领导人员和类型Ⅲ的议员有同样的偏好，那么他就会在议程上提议方案*A*和方案*B*，尽管方案*A*在第一轮投票中获胜，但在步骤3中输给方案*C*。

所以，无论参议院领导人员的偏好如何，通过操作议程，利用子博弈精炼纳什均衡，都会使自己最偏爱的方案通过。**不管他最偏爱哪种方案，他都应该首选其他两种方案，然后使获胜方案和他偏爱的方案竞争，策略导致的结果永远是后一种方案获胜。**注意，这种方法同样适用于与类型Ⅰ议员有同样偏好的参议院领导人员。你只需见证操控议程的力量。

9.3 承诺

在日常生活中，我们不轻易做出承诺。承诺意味着限制自己在未来的选择，理性的做法是推迟，或者完全避免做出承诺。然而，这种逻辑并不总是适合各种战略场景。**承诺可以是有益的，因为它限制了你在未来的行为。益处并非来源于它如何影响你的行为，而是来源于他人在了解你以某种方式限制自己行为后，如何做出回应。**

在这个小节中，我们将在两个商业场景中衡量承诺的价值。第一个例子和许多成功的企业家面临的状况相关：现在假设他的企业已经成功，他如何做才能使别人不模仿他的行为来窃取他的盈利模式？第二个例子和史上的一个谜相关：17世纪时，在和东印度的贸易中，为什么荷兰商人比英国商人更积极主动？正如我们将要讨论的，这不是因为航船或技术的原因，而是因为合同限制了荷兰商人必须更主动。

9.3.1 阻止进入市场博弈

思考一个在市场中盈利运营的在位公司。也许这是一个新的市场，这个公司是第一个进入市场的公司，这家公司的挑战是如何阻止其他竞争者的进入。我们现在分析承诺如何使这个公司达成自己的目标。

如果这个公司打算长期垄断市场，我们假设它的期望盈利（或收益）为1 000。目前有一个新公司想要加入这个有潜力的市场，如果它选择进入，将会面临如图9—19描述的竞争环境。为了使问题更简单，每个公司在3个水平上定价：低价、中价或高价。

		新公司		
		低价	中价	高价
在位公司	低价	300,300	350,325	400,250
	中价	325,350	400,400	500,325
	高价	250,400	325,500	450,450

图9—19　进入市场后博弈（总盈利）

纳什均衡预测到两个公司都将选择中价（实际上，这是一个占优策略），每个公司盈利 400。

在决定是否进入市场之前，潜在进入公司不仅要考虑它将产生的期望收益（刚才已证明为 400）还要考虑进入的成本。假设进入市场的成本为 350，350 可能是建立厂房的成本，也可能是设立店的成本。于是，潜在进入公司可以期望得到纯利 50。假设这个公司不进入市场，得到的收益为 0，那么最佳策略就是进入市场。因为，如果它期望两个公司的定价都符合纳什均衡，它会发现进入市场可以盈利。

对这个扩展式博弈的描述如图 9—20 所示，新公司首先决定是否进入市场。如果它选择不进入，那么在位公司得到垄断利益 1 000，新公司得到的收益为 0。相反，如果新公司选择进入市场，那么两个公司同时做出定价决定。注意图 9—20 中新公司得到的收益是图 9—19 中收益值减去成本 350。

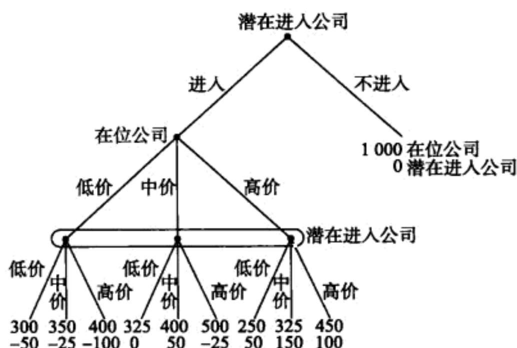


图 9—20　进入市场博弈

在描述公司如何做出决定时，我们利用逆向归纳法来推导。也就是说，我们推出随着新公司进入市场所出现的子博弈的纳什均衡。这个子博弈使得两个公司都选择定价为中价，然后，用纳什均衡替换子博弈：在位公司的收益为 400，新公司的收益为 50，这就使得博弈成为只包含新公司 1 个参与者的博弈，在博弈中，新公司决定是否进入市场，存在一个纳什均衡：进入市场，所以，子博弈精炼纳什均衡是（进入 / 中价，中价）。

对在位公司来说，新公司的进入是最不利的消息，因为相比原来的收益 1 000 来说，现在得到的收益只有 400，所以它想方设法阻止新公司进入市场。有趣的是，存在一个纳什均衡阻止新公司进入市场：策略组合（不进入 / 中价，低价）。给定这个策略规定新公司不进入，不论在位公司的策略是什么，得到的收益都为 1 000，因为如果新公司进入，在位公司才调价。所以，在进入市场后的博弈中，在位公司定低价是最优策略。潜在进入公司实际进入市场时，假定在位公司设定低价，那么如果潜在进入公司进入市场，得到的最高收益为−25（策略规定潜在进入公司定中价）。因为这样得到的收益低于不进入市场得到的收益，所以对潜在进入公司来说，（不进入 / 中价）或（不进入 / 低价）与（不进入 / 高价）是最优策略。

所以，在位公司以低价威胁阻止新公司进入是一个纳什均衡。然而，这种威胁不可信，潜在进入公司也不应该相信这种威胁。换句话说，尽管（不进入 / 中价，低价）是纳什均衡，但不是子博弈精炼纳什均衡。这个子博弈的子策略组合（中价，低价）本身不是一个纳什均衡，只有那些差劲的博弈理论家才会建议在位公司通过设定低价来阻止潜在进入公司进入市场。

跳出图 9—20 中描述的模型，我们提出问题：存在阻止潜在进入公司进入市场的方法吗？我们可以从纽约垃圾拖运公司的运营中得到启示，垃圾拖运公司被有组织的犯罪集团控制，不久另一个公司要进入市场，公司的一名员工在公司邮箱里发现血淋淋的狗头，并附纸条"欢迎来到纽约"。但我们还是思考更合法、更不暴力的方式来阻止新公司的进入。

如果潜在进入公司进入市场，在位公司如何定价才能击败潜在进入公司呢？如果潜在进入公司可以预料到在位公司实施低价策略，那么它选择不进入市场，因为进入市场得到的毛利 325，小于进入市场的成本。所以，在位公司如何使潜在进入公司相信它将会采取低价策略呢？

好吧，我们假设有一种全新的生产技术（尽管成本很高，为 500），可以大大地降低单位生产成本，比如它需要的投资更少（如更少的能量）。如果在位公司要利用新技术，那么潜在进入公司进入市场所面临的竞争环境如图 9—21 所示，这个竞争环境与图 9—19 中所示的不同，因为图 9—21 竞争环境中的在位公司生产成本降低了（注意图 9—21 中所示的收益中包含投资成本）。新公司得到的毛利和以前一样，因为在位公司的生产成本不直接影响新公司的收益；只有在位公司的定价会直接影响新公司得到的收益。在位公司得到的毛利更高，因

为它的成本降低了。

新公司

		低价	中价	高价
在位公司	低价	475,300	525,325	575,250
	中价	425,350	500,400	600,325
	高价	325,400	400,500	525,450

图 9—21　投资新技术后进入市场博弈

图 9—19（没采用新技术）和图 9—21（采用新技术）最大的不同是在采用降低成本的新技术后，在位公司更倾向于设定低价。若不采用新技术，在位公司的最优策略是设定中价。现在采用新技术后，当新公司设定低价或中价时，在位公司都设定低价（当新公司定高价时，在位公司设定中价）。图 9—21 中所示的进入市场博弈中的纳什均衡使得在位公司设定低价，而新公司设定中价。

图 9—22 描述的扩展式博弈考虑了在位公司投资新技术的可能性。在位公司首先决定是否采用新技术，在决定是否进入市场之前，潜在进入公司观测到在位公司的选择。如果它选择不进入市场，那么在位公司得到的收益 1 000（如果不采用新技术）或 700（如果采用新技术）。如果新公司进入市场，那么两个公司同时竞争如何定价，注意，潜在进入公司得到的收益等于它得到的毛利和进入成本之差（如果它选择进入市场），在位公司得到的收益等于它得到的毛利减去投资新技术的成本（如果它采用了新技术）。

为了求解图 9—22 所示博弈的子博弈精炼纳什均衡，我们首先求解两个公司定价的最后两个子博弈中的纳什均衡。用纳什均衡收益替代子博弈，得出图 9—23。求解图 9—23 中树形结构的最后两个子博弈可以证明：如果在位公司不采用新技术，潜在进入公司进入市场（进入市场和不进入市场得到的收益分别为 50 和 0）；如果在位公司采用新技术，潜在进入公司不进入市场（不进入市场和进入市场得到的收益为 0 与-25）。

再返回在位公司的决定，如果不采用新技术，在位公司期望得到收益 400（因为新公司进入市场后，两个公司都设定中价）；但采用新技术，在位公司得到的收益为 700（因为采用新技术后，新公司选择不进入市场）。

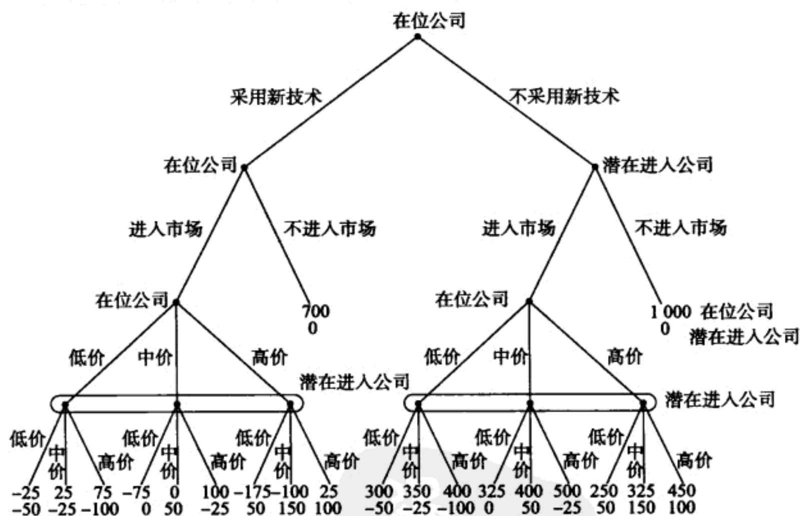


图 9—22　进入市场—投资新技术博弈

总之，唯一的子博弈精炼纳什均衡如下：在位公司采用新技术，如果潜在进入公司随后进入市场，在位公司设定中价。再看潜在进入公司的策略，我们发现如果在位公司投资新技术，那么潜在进入公司选择不进入市场，但如果潜在进入公司做了错误的决定进入市场，那么它设定中价。如果在位公司不采用新技术，那么潜在进入公司进入市场，仍然设定中价。**阻止新公司进入的策略的关键在于分析在位公司设定低价来抵抗新公司进入市场对新公司的影响。**所以，潜在进入公司可以预料到进入市场不盈利的状况。通过改变对低价策略的偏好，

在位公司使得低价策略更具攻击性，采用新技术降低成本使低价策略更占优势。所以，**采用新技术改变了进入市场博弈中在位公司的偏好，这样承诺设定低价阻止了新公司的进入。**和之前的博弈相反，现在假设潜在进入公司在决定是否进入市场之前没有观测到在位公司采用新技术的决定（尽管在进入市场后可以观测到）。博弈结果如图 9—24 所示（其中，我们已经用各自的纳什均衡收益代替了定价子博弈）。现在，对在位公司来说，不采用新技术策略严格优于投资新技术策略，且在位公司不采用新技术，新公司进入市场时，存在一个纳什均衡。那么，如果在位公司投资新技术阻止新公司进入市场，在决定进入市场之前，潜在进入公司首先要确定在位公司已采用了新技术。换句话说，**在位公司承诺要设定低价阻止新公司进入，潜在进入公司应意识到这种承诺的重要性。**

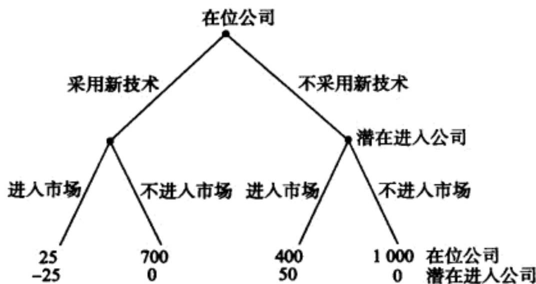


图 9—23　求解最后两个子博弈

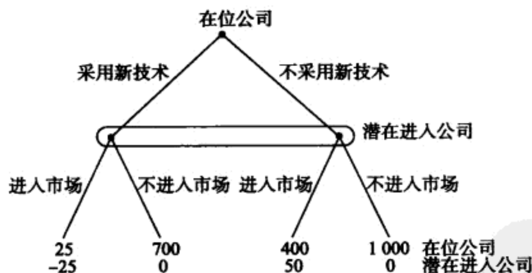


图 9—24　当潜在进入公司对在位公司投资新技术的决定不了解时，进入市场与采用新技术博弈

1964 年上映的传世经典影片《奇爱博士》（*Dr. Strangelove*，又译为《我如何学会停止恐惧爱上炸弹》）体现了承诺行为是共同知识的重要性。一架美国轰炸机正向苏联投放炸药，且轰炸机与外界断了联系。不幸的是，美国了解到苏联有一台"世界末日装置"，在核武器攻击后，"世界末日装置"自动引爆：

> 美国总统墨尔金·马夫雷："世界末日装置"？是什么装置？
>
> 苏联驻美大使得萨斯基：它是一种能毁灭地球上所有动植物生命的终极武器。
>
> 美国总统马夫雷：恐怕我还没有明白。如果我们的战机实施攻击计划，苏联总理威胁要引爆装置吗？
>
> 苏联驻美大使得萨斯基：总统先生，不是这样的，这不是一个理性的人做出的行为。"世界末日装置"被设置为自动引爆。
>
> 美国总统马夫雷：但我认为你们一定可以解除引爆装置。
>
> 苏联驻美大使得萨斯基：不能。任何试图阻止其引爆的行为都会导致其自动引爆。
>
> 美国总统马夫雷：但怎么可能存在一种可以自动引爆，同时又不能解除爆炸的装置呢？
>
> 奇爱博士：总统先生，不是可能存在，而是一定存在。这就是装置的核心技术所在，制造装置就是为了阻止敌人进攻……利用对于袭击的恐惧。因为这个装置是自动的且不可撤销的，这就完全排除了人为的干涉。"世界末日装置"很可怕，也很容易理解，同时也是完全可以令人信服的。（转向得萨斯基）但如果你把"世界末日装置"作为一个秘密，"世界末日装置"就失去了存在的意义！你为什么不向全世界宣布这个消息呢？
>
> 苏联驻美大使得萨斯基在星期一的党会上会宣布。你知道，总理喜欢惊喜。

9.3.2　管理合约与竞争：17 世纪东印度贸易

在 17 世纪，从东印度进口香料和丝绸可以得到丰厚的利润。即使运输成本很高（大约 7% 的航船都被海盗掠夺），但边际利益也很高，因为这些商品将在欧洲以成本 5 倍的价格出售。那时，贸易被英国和荷兰垄断，考虑到他们的航海技术在当时的普遍优越性，这点并不奇怪。而令人困惑的是为什么荷兰比英国在贸易上更主动呢？例如，1922 年荷兰运输了 2 280 000 磅胡椒粉，

超过英国的运输量 1 615 000 磅约 40%。

一个有趣的假设是合约的不同——英国和荷兰商人的工作报酬如何计算。争论在于荷兰商人得到的报酬驱使他们在东印度贸易线上发送更多的货船，且因为他们更主动积极，所以英国商人发送的货船越来越少。为了分析这个假设背后的争论，我们提出了一个简单的模型仅供思考。



假设英国贸易公司和荷兰贸易公司的持股人都只关注公司的利润。然而，一个贸易公司并不直接由持股人经营，而是由经理经营，这决定了经理的行为，也决定了他的报酬。如果他的报酬是公司利润的份额，那么他将竭尽全力最大化公司的利润，因为公司的利润越多，他的薪酬就越高。有趣的是，在这个战略场景中，持股人在追求最高利润时却使得经理一点也不在乎公司的利润！

在这场博弈中，假设仅有两家航运公司，一个是英国的，另一个是荷兰的。每个公司中，持股人决定经理的薪酬合约拟定。假设存在两种可行的合约：（1）经理的薪酬等于公司利润的一部分；（2）经理的薪酬等于公司收入的一部分，收入（是公司卖出商品得到的钱）等于出售商品的磅数乘以每磅的价钱。利润等于收入减去成本，这里的成本是指公司把商品运输到欧洲所付的钱，包括在印度购买商品的钱和运输费。而持股人得到的收益是公司的利润，所以持股人最关心的是如何使利润最大化。

现在进一步假设一旦经理拿到合约（假设他们也同意合约的内容），他们就了解双方的合约。尽管每个经理都知道另一个经理的合约是利润的提成还是收入的提成，他们却不知道另一个经理的薪酬具体是多少。那时，每个经理决定公司承接的航次就决定了商品比如胡椒粉的运输量，为了简化分析过程，假设仅有 3 种运输方式：航次少，航次中，航次多，表 9—2 列出了九种可能策略组合中，各个经理得到的利润和收入。比如，如果英国选择航次中的运输量，荷兰选择航次多的运输量，那么英国公司的利润为 60 000，收入为 112 500，而荷兰公司的利润为 67 500，收入为 135 000。

表 9—2　东印度贸易公司的利润和收入

英国	荷兰	英国公司的利润	英国公司的收入	荷兰公司的利润	荷兰公司的收入
少	少	80 000	120 000	80 000	120 000
少	中	70 000	110 000	87 500	137 500
少	多	62 500	100 000	85 000	150 000
中	少	87 500	137 500	70 000	110 000
中	中	75 000	125 000	75 000	125 000
中	多	60 000	112 500	67 500	135 000
多	少	85 000	150 000	62 500	100 000
多	中	67 500	135 000	60 000	112 500
多	多	57 500	120 000	57 500	120 000

现在假设英国和荷兰公司的经理的薪酬都是利润的一部分。具体来说，薪酬等于利润的 1%。那么经理参与的博弈如图 9—25 所示，这里假设他们得到的收益等于他们的薪酬。[①]这个博弈存在一个均衡，在均衡点上，两个经理都选择航次为中的策略。每个经理得到的收益为 750，每个贸易公司的利润为 75 000，参考表 9—2。

荷兰公司经理（关注公司利润的最大化）

英国公司经理（关注公司利润的最大化）		航次为少	航次为中	航次为多
	航次为少	800,800	700,875	625,850
	航次为中	875,700	750,750	600,675
	航次为多	850,625	675,600	575,575

图 9—25　当每个经理的收益等于公司利润 1% 时的东印度贸易博弈

现在假设荷兰公司经理的收益为公司收入的一部分——具体来说是公司收入的 0.5%。英国公司的经理得到的收益还是等于公司利润的 1%。所以，荷兰公司的经理就忽略购买和运输商品的成本，他通过控制航次

① 收益等于表9—2中公司利润的1%。

的多少来从出售进口商品中获取最高收入。两个公司的经理所参与的博弈如图 9—26 所示。这个博弈存在一个纳什均衡，英国公司经理选择航次少，荷兰公司经理选择航次多。于是，英国公司得到的利润为 62 500（收入为 100 000），荷兰公司得到利润 85 000（收入为 150 000）。

英国公司经理（关注公司利润的最大化） \ 荷兰公司经理（关注公司收入的最大化）		航次为少	航次为中	航次为多
	航次为少	800,600	700,687.5	625,750
	航次为中	875,550	750,625	600,675
	航次为多	850,500	675,562.5	575,600

图 9—26　当荷兰公司经理得到的收益为公司收入的 0.5%，英国公司经理得到的收益为公司利润的 1% 时的东印度贸易博弈

这里有两点需要注意：第一，荷兰公司不仅比英国公司得到的收入多，创造的利润也更多，即使荷兰公司更注重收入的高低，而不是利润的多少。并且，当荷兰经理关注公司的利润最大化时，荷兰公司得到的收入也更高。对图 9—25 中的博弈来说，当荷兰经理只关注公司的利润时（因为他的薪酬是公司利润的一部分），荷兰公司赢得利润 75 000，但现在当公司经理关注公司的收入时（因为他的薪酬是公司收入的一部分），荷兰公司赢得利润 85 000。所以，当经理进口更多的商品，最大化公司收入而不是利润时，持股人得到的利润更多。这种情况是如何发生的？

导致这种结果的关键，是英国贸易公司对荷兰经理关注收入（而不是利润）的最大化会采取什么措施。英国公司是否会继续采取航次为中的策略，从而使荷兰经理的薪酬与收入挂钩，这样荷兰经理将进口更多的商品，降低公司的利润。假如英国经理选择航次为中的策略，当在乎公司收入的荷兰经理选择航次为多而非航次为中的策略时，荷兰公司的利润从 75 000 降到 67 500。但幸运的是，对荷兰公司来说，英国公司了解到荷兰经理将进口更多的商品，所以英国公司不会在市场投放更多的商品以免影响供求关系，从而降低商品的价格。这就是为什么新的纳什均衡使英国公司进口少量的商品的原因，特别是荷兰公司在进口大量商品的时候。进口大量商品的承诺（通过使荷兰经理的薪酬与进口商品量挂钩），使得英国公司进口更少的商品，这样就增加了荷兰贸易公司的利润。


点拨

在一个策略集合中，一个参与者可以通过承诺自己在未来的行为得到更多的收益，如果他把这个承诺泄漏给其他参与者。

本章总结与练习题

本章总结

在求解一个博弈时，关键是寻找最令人信服的策略组合：我们认为聪明的参与者采取何种策略。如在第8章中提到的，纳什均衡并不总是最令人信服的策略，尽管给定其他参与者采取的策略，每个参与者都采取最优策略。一个参与者采取策略的优劣性取决于这样的信念：即如果一个参与者改变了策略，那么另一个参与者会采取非理性行为。承诺未来做出非理性行为被称为不可置信的威胁，且纳什均衡预测不可置信的威胁不具有信服力，所以其他参与者不会相信这种威胁。

对第8章提到的完美信息博弈来说，更严格概念的子博弈精炼纳什均衡剔除了涉及不可置信的威胁的纳什均衡。在这章，我们为完美信息扩展式博弈或不完美信息扩展式博弈的子博弈精炼纳什均衡给出了定义。

实施子博弈精炼纳什均衡分两步：第一步，你必须确定博弈中的博弈也是一个完整的博弈。**“博弈中的博弈”**被称为**子博弈**。第二步，我们定义了子策略组合是策略组合的一部分，它与子博弈相关。然后，使子策略组合形成子博弈的纳什均衡。所以，子博弈精炼纳什均衡规定参与者在每个子博弈中都采

取最优策略，这样，当一个参与者观测到对手做出非理性行为时，他应该认为所有参与者都会将自己利益最大化，从而不相信这种威胁。

在博弈理论中，**承诺**是指在未来做出某种行为，可以是做出投资行为来影响未来行动的收益，也可以是限定自己在未来的选择。在策略集中，承诺是有价值的，因为它可以影响其他参与者的策略。一个众所周知的承诺可以束缚一个参与者未来的行为，从而诱导其他参与者调整自己的策略，而这或许正是做出承诺的参与者所期待的结果。理解承诺的价值是博弈论中的一个重要课题。

练习题

1. 为图 9—27 中的博弈推导所有的子博弈精炼纳什均衡。

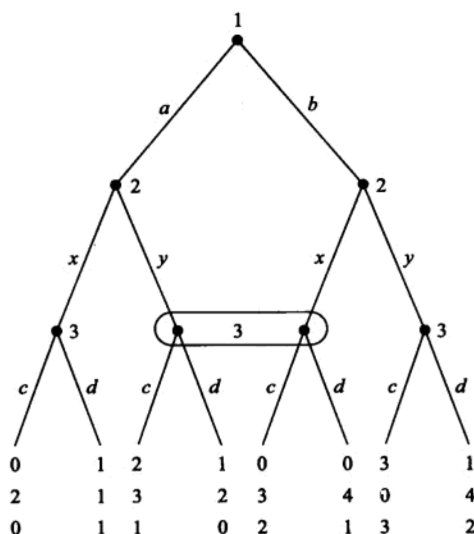


图 9—27

2. 在图 9—28 的 OS/2 博弈变体形式中，信息集的结构与原博弈不同，参与者得到的收益也不同。现在如果 IBM 开发 OS/2，软件公司 1 先行决定是否开发兼容软件，其次，在公司 1 的选择被发布后，公司 2 和公司 3 同时采取策略。

a. 推导所有的子博弈纳什均衡；

b. 推导一个不是子博弈精炼纳什均衡的纳什均衡，并解释它为什么不是子博弈精炼纳什均衡。

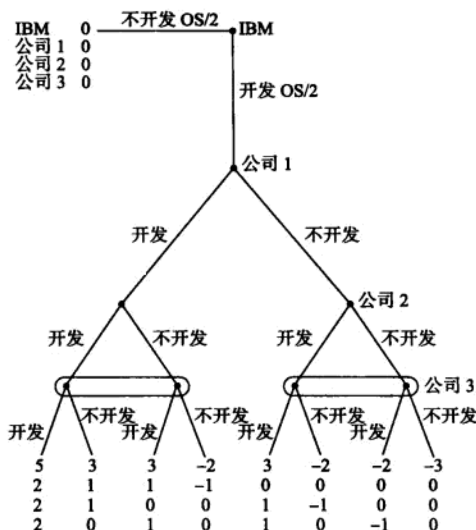


图 9—28 OS/2 博弈的变体形式

3. "我不能和你一起走"赫敏脸红着说道，"因为我已经答应和其他人一起去。"

"不，你没有。"罗斯说道："你刚刚说你已经摆脱内维尔了。""哦，我说过吗？"赫敏说。她的眼睛闪烁着不安，"这仅仅因为你关注了我三年吗？罗斯，这并不意味着别人没有注意到我是个女孩。"

在圣诞舞会的前一周，维克多和罗斯都在考虑是否邀请赫敏。如图 9—29 所示，维克多先行决定是否邀约赫敏（注意，邀一个女孩子去舞会比与流氓斗殴更可怕）。如果他鼓起勇气邀请她，那么赫敏决定是否接受邀请。在维克多邀请后，罗斯最终决定克服紧张心理，

告诉赫敏他对她的感情（也邀请她去舞会）。然而，注意在这个信息集中，罗斯不知道维克多与赫敏之间发生了什么。罗斯不知道维克多是否邀请了赫敏，如果维克多邀请了，赫敏是否接受了邀请。如果罗斯邀请了赫敏，且赫敏没有打算和维克多一起去舞会（或许因为维克多没有邀请她，或许因为她拒绝了维克多），那么赫敏必须决定是否答应罗斯的邀请。在树形结构中赫敏的决策节点上，她知道自己参与的博弈环境，因为她完全了解所泄露的全部信息。考虑到得到的收益情况，赫敏宁愿和罗斯一起去舞会而不是和维克多一起参加舞会。罗斯和维克多都想同赫敏一起参加舞会，但如果她不能去或不愿意去，那么他们宁可不发出邀请。利用子博弈精炼纳什均衡的概念分析接下来发生的情况。

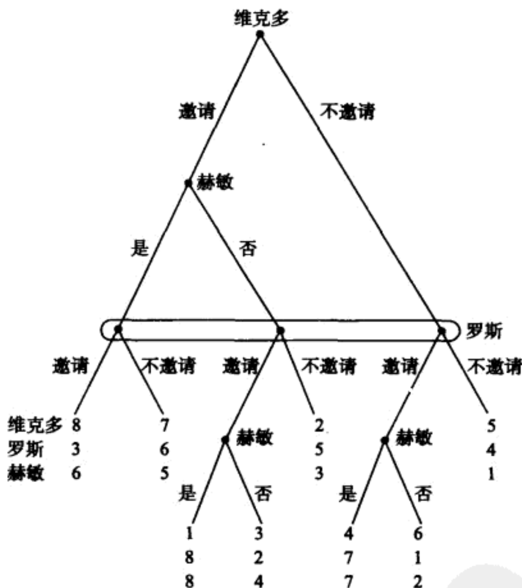


图 9—29　圣诞节舞会

4. 以练习题 3 中的圣诞舞会为例，现在假设在决定是否邀请赫敏之前，罗斯观察到维克多是否邀请了赫敏。然而，如果维克多确实邀请了赫敏，当罗斯决定是否邀请赫敏时，他不知道赫敏是否接受了维克多的邀请。

a. 写出扩展式博弈。

b. 推导出所有的子博弈精炼纳什均衡。

5. 在一次战争中，A 国研制了一艘有无声喷水推进系统的核潜艇，因而可以逃过对方的探测器。在执行首次航行时，舰长决定背叛，因为他认为这项科技会破坏交战两国军事力量的平衡，引发更大规模的战争。他集合了一些忠诚于他且同意背叛的船员，然而，舰长担心船员可能会在航行中改变主意，并且如果一个船员认为其他船员改变主意，他也会改变主意。于是，舰长想写一封信（在核潜艇出发后寄送给 A 国政府）陈述他背叛的计划。这个博弈的扩展式如图 9—30 所示。

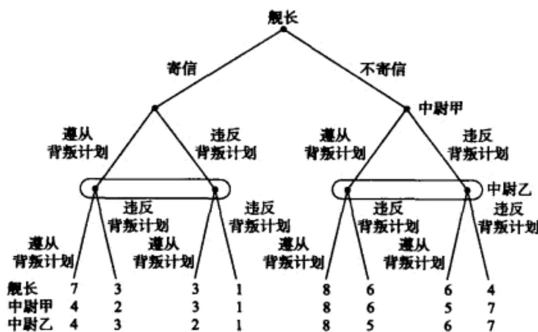


图 9—30　核潜艇舰长背叛

舰长首先决定是否寄信。在把他的决定透露给船员后（在他们都出海后），为保险起见，他又把计划透露给他的两个亲信中尉甲和中尉乙。之后，他的船员们被控制起来，于是船员们开始决定是遵从背叛计划还是违反背叛计划。收益情况表明这三个参与者都选择背叛，且认为不应寄信给自己的政府（后来 A 国政府派出另一艘核潜艇击沉了这艘潜艇）。

a. 推导所有的子博弈精炼纳什均衡。

b. 解释说明舰长为什么选择寄信。

6. 让我们返回并分析东印度贸易的竞争环境。

a. 思考当英国公司和荷兰公司的经理薪酬都是公司收入的 0.5% 时的博弈，博弈收益如图 9—31 所示。求解所有的纳什均衡。

荷兰公司经理（关注公司收入最大化）

英国公司经理（关注公司收入最大化）		航次为少	航次为中	航次为多
	航次为少	600,600	550,687.5	500,750
	航次为中	687.5,550	625,625,	526.5,675
	航次为多	750,500	675,562.5	600,600

图 9—31　经过修改的东印度贸易博弈

b. 现在思考英国公司和荷兰公司持股人和他们的经理之间的合约博弈。假设他们同时选择，一个持股人选择经理薪酬为公司利润的 1%，另一个选择经理薪酬的为公司收入的 0.5%。和之前的假设相同，持股人的收益等于公司的利润（我们忽略他们付给经理的少量薪酬）。在选择两个合约后，两个合约的内容都透露给经理，然后经理选择航次为少，航次为中，航次为多的策略。求解所有的子博弈精炼纳什均衡。

7. 思考图 9—32 中的扩展式博弈。终端节点首行收益是参与者 1 得到的收益。求出所有的子博弈精炼纳什均衡。

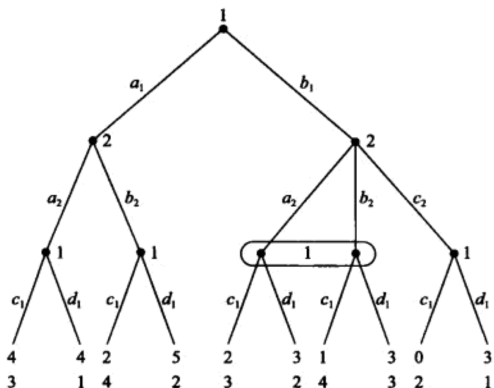


图 9—32

8. 思考图 9—33 中描述的扩展式博弈。终端节点首行的收益是参与者 1 得到的收益，中间行是参与者 2 得到的收益，底行是参与者 3 得到的收益。

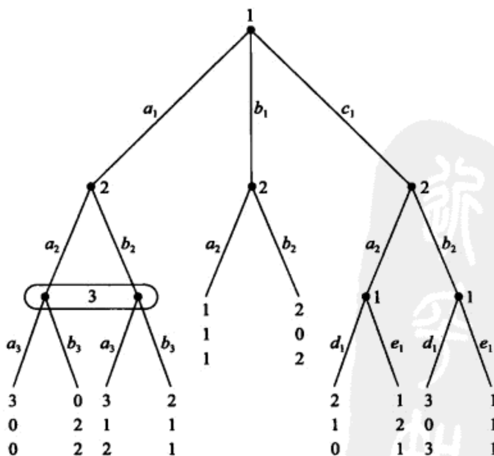


图 9—33

a. 推导每个参与者的策略集。

b. 推导所有的子博弈精炼纳什均衡。

c. 推导一个不是子博弈精炼的纳什均衡，并解释说明为什么不是子博弈精炼纳什均衡。

9. 思考图 9—34 中描述的扩展式博弈。终端节点首行收益是参与者 1 得到的收益，底行收益是参与者 2 得到的收益。

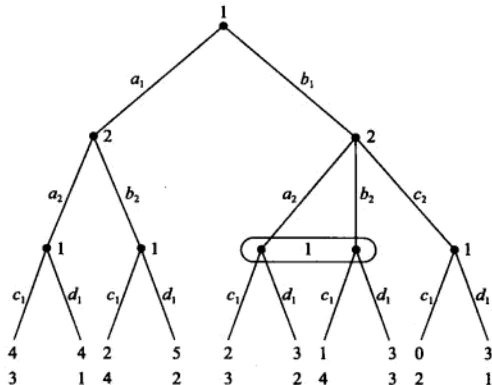


图 9—34

a. 描述每个参与者策略的一般表达式。

b. 推导所有的子博弈精炼纳什均衡。

10. 思考图 9—35 中的扩展式博弈。终端节点的第一行数字是参与者 1 得到的收益，第二行数字是参与者 2 得到的收益，第三行数字式参与者 3 得到的收益，最后一行是参与者 4 得到的收益。

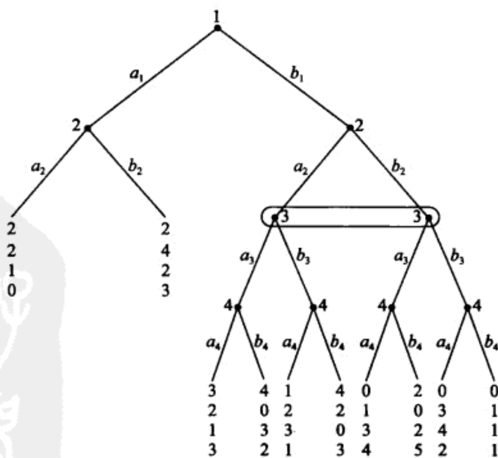


图 9—35

PDG

a. 推导每个参与者的策略集。

b. 推导所有的子博弈精炼纳什均衡。

11. 棒球中一个最有意思的现象是“抢分触击”。这种情况涉及的球员是三垒的一个跑垒员，以及少于两个跑垒员出局的情况。当投球手挥臂准备向本垒投球时，这个跑垒员快速冲向本垒。击球手的任务是盯住这个球并短打球远离本垒，所以没有对手有机会防守这个球从而将跑垒员触杀出局。另一组队员可以实施“投手故意投坏球”来阻挠这个战术的使用。“投手故意投坏球”使得击球手无法击球。接球手知道球正飞过来，腾空跨越接住球，这样很容易将三垒的跑垒员触杀出局。当然，这个队的教练可能意识到对方会设计“投手故意投坏球”，并取消“抢分触击”。如果另一个教练确实实施了“投手故意投坏球”，但这个教练没有实施“抢分触击”，那么他使投手处在一个更复杂的情境。



棒球中的“抢分触击”

让我们思考两个优秀且懂得如何使用战术的教练（托尼·拉鲁萨和乔·托尔）所参与的博弈情况。他们所参与的博弈如图 9—36 所示。为了更简化模型，我们假设击球手的当前击球数是两个坏球和一个好球，所以托尔（他的对员正投球）可以实施至多一个“投手故意投坏球”。比赛开始时，两个教练同时采取行动：托尔决定是否实施“投手故意投坏球”，拉鲁萨决定是否使用“抢分触击”。如果拉鲁萨使用了“抢分触击”，托尔实施了“投手故意投坏球”，那么对托尔来说，他获得最佳结果，得到的收益为 10；而对拉鲁萨来说，获得最坏结果，得到的收益为 0。如果托尔实施了“投手故意投坏球”，而拉鲁萨没有使用“抢分触击”，那么拉鲁萨得到的收益为 8，托尔得到的收益为 2。如果拉鲁萨使用了“抢分触击”，而托尔没有实施“投手故意投坏球”，那么这个结果正是拉鲁萨期待的结果，他得到的收益为 9，托尔得到的收益为 1。最后，如果拉鲁萨没有使用“抢分触击”，托尔没有实施“投手故意投坏球”，那么假设策略情境转移到下一个投球，这时托尔再次决定是否实施“投手故意投坏球”，拉鲁萨决定是否使用“抢分触击”。在混合策略中求解一个子博弈精炼纳什均衡。

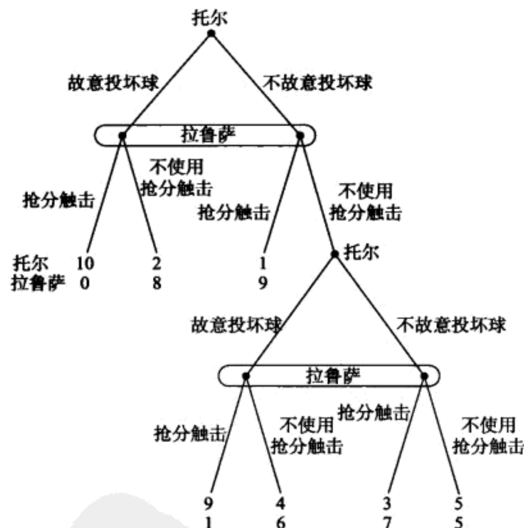


图 9—36　棒球中的“抢分触击”

第四部分 不完全信息博弈

Games, Strategies, and Decision Making

第10章 我知道你所不知道的信息：拥有私人信息时的博弈

学习目标

10.1 不完全信息博弈：《慕尼黑协定》
10.2 贝叶斯博弈和贝叶斯纳什均衡
10.3 当所有的参与者都拥有私人信息时：拍卖博弈
10.4 委员会和陪审团的投票

本章导读

谁会料到有什么邪恶潜伏在人类的心灵中？

——*《魅影奇侠》（The Shadow）*

把策略情境模拟成一个博弈模型最重要的一个假设是：对所有的参与者来说，博弈是共同知识。每个参与者知道谁参加了博弈、他们各自的策略选择以及每个参与者如何评估结果。对于一些情况而言（比如早上起床你和你的姐妹们争着去洗手间洗漱）这个描述就很贴切，而在其他情况中，共同知识很明显是违反常识的，而这些知识的缺乏对于参与者的行动有很大影响。比如，假设你走进一家汽车经销商与卖家谈判汽车的价格，你应该出的价钱不仅取决于你所认为汽车的价值，而且取决于经销商愿意以多少价钱出售汽车。问题在于你不知道经销商的最低卖价，经销商也不知道你愿意出的最高买价。这就是拥有私人信息时的情况。

在现实生活中有许多拥有私人信息的博弈。模拟和求解这些策略情境是极其困难的，直到20世纪60年代后期，一个名为约翰·海萨尼（John Harsanyi）的匈牙利天才数学家在这方面做出重大突破后才解决了这方面的难题。他提出了一种将一个具有不完全信息的博弈转换成一个具有完全（但不完美）信息博弈的方法。通过这种转换方法，不完全信息博弈被转换成一个等价的完全信息博弈，从而可以对原来的不完全信息博弈进行研究。听起来很神奇，不是吗？然而，一旦解释清楚，这些道理也是浅显易懂的。不过这些道理并不会因为浅显易懂价值下降。

10.1 不完全信息博弈：《慕尼黑协定》

在海萨尼因为对博弈理论做出巨大贡献而获得诺贝尔经济学奖之前，他已经展示了个人的博学和勇气。因为是匈牙利的犹太人，他在1944年从一个火车站开始逃亡生涯，以躲避纳粹集中营的迫害。那时，匈牙利政府变得越来越专制和暴力，在1950年他穿过一个戒备相对较松的沼泽地逃离匈牙利，移民海外。鉴于他个人如此大的成就，我们就利用他首创的方法来分析第二次世界大战前一个重要的外交事件。

让我们把时钟拨回到1938年，德国纳粹组织强占了奥地利，且已证实阿道夫·希特勒正在酝酿对捷克斯洛伐克的苏台德区采取相同的行动。因为第一次世界大战的硝烟才刚散去不久，欧洲人害怕再次经历这类痛苦和恐怖的战争。为了维持和平，英国首相内维尔·张伯伦（Neville Chamberlain）去德国与希特勒达成协议。1939年9月28日，张伯伦和希特勒签署《慕尼黑协定》（*Munich Agreement*），同意希特勒占领苏台德区来换取希特勒不再进攻其他地区的"保证"。这个退让被当做是延迟战争爆发的权宜之计。结果证明和平依然是水中月，镜中花。因为德国次年春天就进驻布拉格、侵入波兰，挑起了第二次世界大战。

在决定是否提出和随即签订这份协定时，张伯伦对希特勒的最终意图并不确定。希特勒仅仅是想为德国人民谋取更多的生活空间吗？如果情况是这样的话，那么诸如苏台德区的割让或许会安抚希特勒并避免战争的爆发。亦或是希特勒在编织一个更大的阴谋来侵略欧洲呢？

图10—1很好地诠释了慕尼黑这个例子。张伯伦首先决定是退让还是坚守。假设希特勒会接受退让协定，我们只需关注关于发动战争的决定。张伯伦的选择很明显：他的最优结果是坚守立场避免战争，他的最劣结果就是提出退让，但希特勒还是发动战争。因为提出退让策略，这就使得希特勒有更多的时间去准备武器备战；所以，我们必须假设张伯伦发现这样的结果劣于坚守的结果，从而也参与战争。

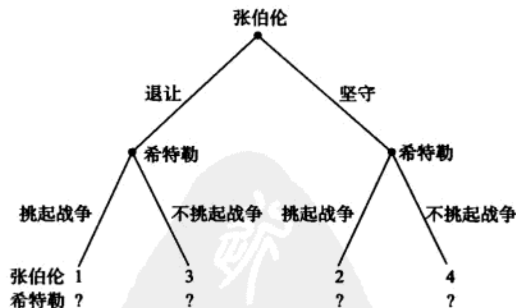


图10—1　不知道希特勒收益函数的情况下的慕尼黑协定博弈

分析这场博弈的难点在于计算希特勒的收益。希特勒知道自己得到的收益是多少，而张伯伦不知道。不知道希特勒的收益，张伯伦如何确定希特勒将会采取什么行动？



希特勒（右）与张伯伦（左）在慕尼黑会见

在我们思考这个难题的解决方法时，不妨先考虑张伯伦思想斗争的各种可能。他的一种想法是希特勒是随和的，图10—2a表示各自得到的收益。我们称希特勒随和是因为他的最优结果是同意退让且不挑起战争。注意，如果张伯伦立场坚定，希特勒就会发动战争掠夺更多的土地。所以，如果张伯伦面对的真的是一个随和的对手且他自己知道这一事实，那么他就应该提出退让策略。



(a) 当希特勒随和时的慕尼黑协定博弈



(b) 当希特勒好斗时的慕尼黑协定博弈

图10—2　慕尼黑协定博弈的两种情况

另一种可能是希特勒是好斗的情况，图 10—2b 总计出各自的收益。这里，希特勒的最优策略是挑起战争，尽管他在接受退让条件之后仍然选择这样做。对于张伯伦来说，他最好还是坚守立场。

实际上，张伯伦不确定自己参与的是图 10—2a 中描述的博弈还是图 10—2b 中描述的博弈。这种类型的博弈就是不完全信息博弈。记得在不完美信息博弈中，参与者知道他们所参与的博弈类型，但在博弈中的某个阶段，一个参与者不知道自己在博弈中所处的位置，也就是说他不知道其他某些参与者过去所采取的行动。相反，在不完全信息博弈中，参与者并不知道他们所参与的博弈类型，博弈的有些信息不是共同知识。尽管这类博弈有多种变体形式，本章主要讨论信息缺失如何影响其他参与者的收益。

求解不完全信息博弈的技巧就是把它转换为不完美信息博弈——就是说，把我们不知道如何解决的东西转换成我们知道如何解决的东西！这样做需要引入一个被称为“自然”的新的参与者。“自然”是一个虚拟的参与者，并不指代树、跳蚤或是蜜蜂，而是指参与者博弈环境中的随机力量。“自然”以外在的形式规定了采取各种行动的概率，呈现参与者关于随机事件的信念。在接下来的文章中，“自然”决定希特勒的最优策略（或收益），如图 10—3 所示。

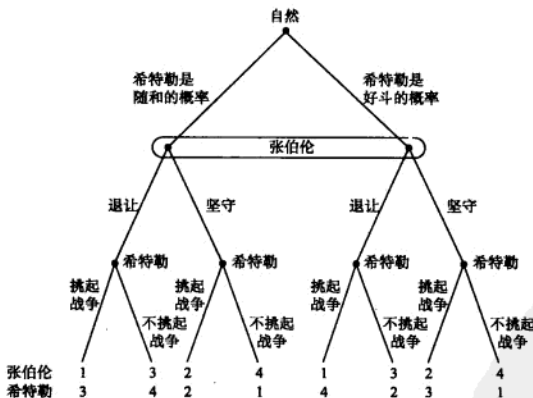


图 10—3 “自然”决定希特勒是随和还是好斗时的慕尼黑协定博弈

在这个模型中，“自然”先选择希特勒是随和的还是好斗的。“自然”做出的这步行动张伯伦是观察不到的——这样他对希特勒的收益缺乏了解；但希特勒是知道这个收益的——因为他知道自己的选择。“自然”赋予这两种可能的概率是共同知识这一假设是很重要的，这里假设希特勒是随和的概率为 60%，是好斗的概率为 40%。

图 10—3 中的扩展式博弈清晰地定义了不完美信息博弈。这个博弈是共同知识，尽管当张伯伦采取行动时不知道“自然”是如何选择的。为了求解纳什均衡，首先注意希特勒的策略包含四种行动，因为他有四个信息集描述“自然”如何选择和张伯伦如何采取行动：（随和，退让），（随和，坚守），（好斗，退让），（好斗，坚守）。于是希特勒的策略就表现为以下四种形式：

如果我是随和的且张伯伦提出退让策略，那么 ______ [填入挑起战争或不挑起战争]。

如果我是随和的且张伯伦提出坚守策略，那么 ______ [填入挑起战争或不挑起战争]。

如果我是好斗的且张伯伦提出退让策略，那么 ______ [填入挑起战争或不挑起战争]。

如果我是好斗的且张伯伦提出坚守策略，那么 ______ [填入挑起战争或不挑起战争]。

这就意味着希特勒的策略集包含 16 种策略，在四处空白处，每处都可填入两种可行的行动。

我们可以直接推导出希特勒的最佳策略。如果他是随和的（如“自然”选择的那样），那么当张伯伦提出退让时，他应该选择不挑起战争；当张伯伦提出坚守时，他应该选择挑起战争。如果希特勒是好斗的，那么不管张伯伦提出什么，他都应该挑起战争。所以，对希特勒来说，最佳策略是（不挑起战争，挑起战争，挑起战争，挑起战争）。

考虑到希特勒的策略，张伯伦如何采取行动呢？张伯伦的策略是一种单方面的行动，因为他只有一个信息集。因为对希特勒的参数不确定，他不能肯定希特勒对他的行动采取何种策略。所以，张伯伦需要计算出他的两个策略得到的期望收益。

张伯伦从退让中得到的期望收益为：

$$0.6\times3+0.4\times1=2.2$$

假设“自然”选择希特勒是随和的概率为 0.6，那么希特勒的最佳策略决定了——如果张伯伦提出退让，他就不挑起战争，张伯伦在这个博弈中得到的收益为 3。希特勒是好斗的概率为 0.4，根据他的策略，他将挑起战

争——这样张伯伦从退让提议中得到的收益为1。用苏台德区来平息希特勒使得张伯伦的期望收益为2.2。相反，如果张伯伦坚守，那么他得到的收益为：

$$0.6\times2+0.4\times2=2$$

只要他提议坚守，不论希特勒的类型是什么，希特勒都会挑起战争，所以得到的收益为2。那么张伯伦的最佳策略是提议退让，因为退让得到的收益为2.2。总之，我们得出这个博弈的解是：张伯伦提议退让，这种情况下如果希特勒是随和的，他就不挑起战争；如果希特勒是好斗的，他就挑起战争。

10.2 贝叶斯博弈和贝叶斯纳什均衡

在这个小节中，我们将更系统地描述如何求解不完全信息博弈。正如我们在10.1小节中学到的，关键在于把不完全信息博弈转换成不完美信息博弈，即贝叶斯博弈。贝叶斯博弈是标准博弈的一个变体，给标准博弈加了一个首阶段，在这个阶段，“自然”决定参与者拥有的私人信息。构建一个贝叶斯博弈的三个步骤：

步骤1：在一个不完全信息博弈中，一个参与者缺乏与自己互动的对方的相关信息。构建贝叶斯博弈的第一步是详述问题的答案，“这个参与者是谁？”我们将参与者拥有的私人信息（所以其他的参与者不知道）称为参与者的**类型**（type），所有的类型集合就是**类型空间**（type space）。参与者的类型就是他的收益，这种情况下，类型空间就是这个参与者各种可能的收益的集合。在慕尼黑协定博弈中，希特勒的类型空间包括随和的和好斗的两种收益。

步骤2：在规定了类型空间后，下一步就是确定每个参与者的类型。当然，在你见到这个参与者之前，他的类型已经被确定了：当张伯伦在慕尼黑出现时，希特勒或者是随和的，或者是好斗的。然而，在贝叶斯博弈中，我们需要将“时光倒流”至参与者不知道自己类型的时候（在了解私人信息之前）。博弈必须这样开始，即每个参与者都被看做一个白板。在这个阶段，随机力量（“自然”）决定每个参与者的类型并把这些类型“写”在那个白板上。“自然”对每个类型赋予概率，这个概率表示了“自然”为一个参与者选择类型的可能性大小。我们假设对所有的参与者来说，这些概率是共同知识。

步骤3：构建贝叶斯博弈的第三步是定义策略集。记得策略是在博弈开始之前就完全规定了的选择的决定规则。贝叶斯博弈中的策略选定是在“自然”行动之前。因为“自然”给定了参与者的类型与一个参与者可能知道的其他信息（和其他参与者如何行动相关），策略规定了每种情况下参与者会采取的行动，这个策略陈述了该如何去做。这个场景类似于将我们重新投胎：如果时光倒流，我是一匹马、一条蛇或是一只蜣螂时，我将怎么采取行动？（哦，千万不要让我投胎为一只蜣螂！）

你可能还在思考第三步。希特勒是一个特殊的类型，实际上，现在我们知道他是好斗的，那为什么还要假设他是随和的呢？实际上设计这个步骤是为了帮助张伯伦分析如何做出决定。张伯伦必须推算出如果希特勒是随和的会采取什么策略，以及如果希特勒是好斗的会采取什么策略，这样他就可以形成一种关于希特勒将会采取什么行动的信念，从而可以设计自己的最优策略。

在对不完全信息博弈进行适当修改使其转换成不完美信息贝叶斯博弈之后，接下来我们试着求解这个博弈。最常用的解决方法是贝叶斯纳什均衡，它是一种规定参与者的每个类型在给定其他参与者的策略下所采取的最佳行动的策略组合，且对每个参与者来说都是如此。

为了区分贝叶斯纳什均衡和纳什均衡，我们记得纳什均衡是这样一种策略组合，在给定其他参与者策略的条件下，每个参与者都选择自己的最优策略从而使自己的收益最高。这就提出了n个条件，每个参与者一个条件。

现在我们思考贝叶斯博弈，令 m_i 表示参与者 i 的类型个数。那么，要使参与者 i 的策略成为一个贝叶斯纳什均衡，对 m_i 种类型来说，这个策略都必须是最优策略。也就是说，对参与者 i 的每个类型来说，采取的行动都是最佳行动。那么如果我们要得出参与者 i 策略为最佳策略的话，就必须满足 m_i 个条件。

前面的这些都是介绍性地描述了贝叶斯纳什均衡；在本章附录里会更正式地定义贝叶斯纳什均衡。如果你不喜欢那些复杂的符号解释，那我们就举例阐释这个概念。

案例研究

狂野西部枪战

在 1875 年的堪萨斯州道奇城（Dodge）的当地沙龙有这样一个争论。警长马歇尔•怀亚特•厄普（Marshal Wyatt Earp）去维持秩序，一个陌生人拉住他的手，好像要开枪。厄普和陌生人都面临一个决定：是立即开枪还是选择僵持。厄普枪战的实力是大家公知的，但厄普却不知道陌生人的情况。他是一个熟练的枪手还是一个普通的牛仔？

实际上，如果这个陌生人是个熟练的枪手且厄普知道这一点，那么图 10—4a 中描述的是陌生人和厄普之间的博弈结果。因为厄普倾向于避免枪战，他得到的最好结果就是他和陌生人都选择僵持来避免枪战的发生。最坏结果就是陌生人选择枪战而厄普选择僵持。假定这个陌生人是个熟练的枪手，厄普很可能会被杀掉。注意厄普的最佳反应是依据自己对陌生人的判断，并做出和陌生人同样的选择：如果他认为陌生人不会开枪，那么他选择僵持；如果他认为陌生人会开枪，那么他也开枪。相反，陌生人选择最优策略，即选择开枪。如果他认为厄普会开枪，那么他一定会开枪，但即使厄普不开枪，他也会选择开枪，因为这会提高他作为一个熟练枪手的名誉（这显然无法成立，因为厄普根本就不知道他是谁）。

另外，假设这个陌生人只是个普通的牛仔。那么，在完全信息博弈中，两个人的博弈收益矩阵如图 10—4b 所示。在这场博弈中，厄普的最优策略是僵持。因为即使陌生人开枪，如果陌生人的确是个普通的牛仔，厄普相信自己可以先开枪打中牛仔。在这种情况下，他既没挑衅陌生人，也防卫了自己。这个牛仔意识到他的能力远不及厄普的能力，所以只有当他认为厄普会开枪时他才开枪。牛仔的最优策略是他认为厄普怎么做，他才怎么做。

陌生人（熟练的枪手）

		开枪	僵持
厄普	开枪	2,3	3,1
	僵持	1,4	8,2

（a）当陌生人是熟练的枪手时的枪战博弈

陌生人（普通的牛仔）

		开枪	僵持
厄普	开枪	5,2	4,1
	僵持	6,3	8,4

（b）当陌生人是普通的牛仔时的枪战博弈

图 10—4　枪战博弈

这场博弈是不完全信息博弈，因为厄普不知道他参与的博弈是图 10—4a 还是图 10—4b 描述的博弈。不知道博弈对手的情况，厄普就不知道如何采取行动。同样，陌生人也不能轻易决定自己如何采取行动，因为如果他是普通的牛仔，对他来说，最佳的策略就是厄普怎么做，他就怎么做。但厄普不确定对手的情况，所以不知道自己如何采取行动，这样陌生人也就无法判断厄普会采取什么样的行动，也就不确定自己的行动方案。

为了解决这个问题，谢里夫•海萨尼（Sheriff Harsanyi）去镇上转了一圈，把这个不完全信息博弈转换成了不完美信息博弈（见图 10—5）。“自然”先行决定这个陌生人是一个熟练的枪手的概率为 0.75，或是一个普通的牛仔，概率为 0.25。厄普无法观察到“自然”的决定，这样我们就可以推断厄普对陌生人的能力和意图都一无所知。厄普只有一个信息集，所以，他的策略集包含开枪和僵持两种。对陌生人来说，他可以知道“自然”的决定——因为他知道自己是熟练的枪手还是普通的牛仔，但他无法猜测到厄普的行动；厄普和陌生人同时采取行动。陌生人有两个信息集，一个是“自然”决定他是一个熟练的枪手，另一个是普通的牛仔。陌生人的策略集就包含四种策略：（开枪，开枪），（开枪，僵持），（僵持，开枪），（僵持，僵持）。如果他是个熟练的枪手，他就会采取第一种策略；如果他是普通的牛仔，他就会采取第二种策略。

在推导贝叶斯纳什均衡时，注意当陌生人是熟练枪手时，他必须选择开枪，在这种情况下，开枪是最优策略。因为对陌生人来说，我们可以为他剔除策略（僵持，开枪）和（僵持，僵持）。还有其他四种候选策略可能会成为贝叶斯纳什均衡：

A: 厄普选择开枪，陌生人选择（开枪，僵持）。
B: 厄普选择僵持，陌生人选择（开枪，僵持）。
C: 厄普选择开枪，陌生人选择（开枪，开枪）。
D: 厄普选择僵持，陌生人选择（开枪，开枪）。

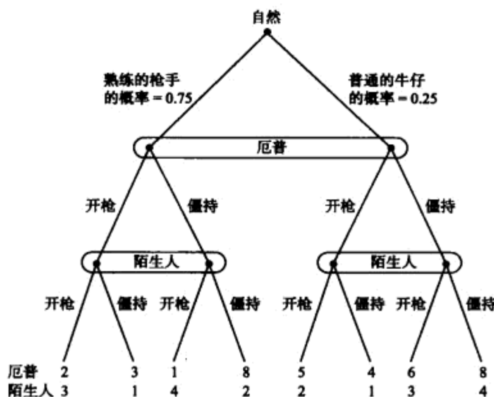


图 10—5　贝叶斯枪手博弈

很明显，策略集合 A 不是一个贝叶斯纳什均衡。如果厄普选择开枪，那么不论陌生人是熟练的枪手还是普通的牛仔，都会选择开枪。然而，只有当陌生人是熟练枪手时，（开枪，僵持）策略才会使陌生人选择开枪。注意策略组合 D 也不是一个贝叶斯纳什均衡：如果厄普选择僵持，那么如果陌生人是普通的牛仔，他最好的策略就是僵持。所以我们就剔除了策略组合 A 和 D。

接下来，思考策略组合 B。在这个策略组合中，厄普选择僵持，如果陌生人是熟练枪手就会开枪，如果陌生人是普通牛仔就会选择僵持。要使策略组合 B 成为一个贝叶斯纳什均衡策略，就必须满足以下三个条件：

1. 厄普选择僵持会使他的期望收益最大化，考虑到陌生人选择开枪的概率为 0.75——因为陌生人是熟练枪手的概率为 0.75，且根据陌生人的策略（开枪，僵持），他最终会选择开枪；陌生人选择僵持的概率为 0.25——因为陌生人是普通牛仔的概率为 0.25，且根据陌生人的策略（开枪，僵持），他最终会选择僵持。
2. 如果陌生人是一个熟练的枪手，那么开枪会最大化他的收益，厄普的策略是选择僵持。
3. 如果陌生人是一个普通的牛仔，那么僵持会最大化他的收益，厄普的策略是选择僵持。

条件 1 保证了厄普的策略是最优策略，条件 2 和条件 3 保证了对两种类型的陌生人来说，采取的策略都是最优策略。首先，注意策略组合 B 肯定满足了条件 2：熟练的枪手会优先选择开枪（不论厄普的策略是什么）。接下来，注意策略组合 B 也满足条件 3：因为当他是一个普通的牛仔时，陌生人一定会选择避免与厄普交锋，那么他会选择僵持，所以他也期望厄普采取僵持的策略。

根据条件 1，厄普采取开枪策略得到的期望收益是：

$$0.75\times2+0.25\times4=2.5$$

陌生人是熟练枪手的概率为 0.75，这样，根据陌生人的策略，他会选择开枪——厄普得到的收益为 2。陌生人是普通牛仔的概率为 0.25，这样，根据陌生人的策略，他会选择僵持——厄普采取开枪策略得到的收益为 4。相反，厄普采取僵持策略得到的收益是：

$$0.75\times1+0.25\times8=2.75$$

这样就满足了条件 3，因为采取僵持策略得到的期望收益大于采取开枪得到的期望收益。

10.1 理解测验

推导策略组合C是不是一个贝叶斯纳什均衡。

在慕尼黑协定博弈中，张伯伦知道关于希特勒的收益情况，希特勒是不需要清楚这一点的。这是因为希特勒可以观测到张伯伦的行动，要使自己的收益最大化，他所需要知道的仅仅是他自己的类型和张伯伦选择的行动。于是，求解这个博弈只需要验证张伯伦关于希特勒类型的信念。而枪战博弈就不同了，因为厄普和陌生人是同时行动且陌生人的最佳行动取决于他对厄普行动的信念，仅仅验证厄普对陌生人类型的信念是不够的。普通的牛仔需要知道厄普的信念，然后他才能猜测厄普的下一步行动。在这种情况下，我们假设赋予陌生人不同类型的概率值为共同知识就很重要。

10.3 当所有的参与者都拥有私人信息时：拍卖博弈

迄今为止，我们已经探索了只有一个参与者拥有私人信息的博弈场景。张伯伦并不知道希特勒的真实偏好，但希特勒知道张伯伦的真实偏好；厄普并不清楚陌生人的能力和意图，而陌生人知道关于厄普的所有信息。然

而，在许多情况下，每个参与者都拥有私人信息。接下来我们就分析这类情况下的博弈。

一个经典的例子就是拍卖博弈。每个投标者知道她自己愿意出多少钱，但不知道其他投标者愿意出的最高价。在决定如何出价之前，投标者必须预测其他投标者如何出价，而这取决于投标者对拍卖品的价值估计，对每个投标者来说，这都是私人信息。我们将这个场景模拟为贝叶斯博弈，并得出关于出价行为的一些启示。

以一级密封价格拍卖为例来开始我们的探索之旅。在一级密封价格拍卖中，所有的投标者同时提交自己的出价，但他们不知道别的投标者的出价，出价最高的投标者获得物品，并按她自己的出价付钱给拍卖商。我们将在两种截然不同的信息集探索这类拍卖博弈。一个信息集是**独立私人价值**（independent priuate values），它假设不同的投标者对物品的价值有不同的估计，他们之间的估计并没有关联。也就是说，每个投标者有自己的估价，如果她赢标了，物品对她的价值就等于她的估价。投标者的类型就是她的估价。这类博弈之所以被称为独立私人价值，是因为每个投标者有自己的估价，且所有投标者的估价都是随机变量。第二个信息集被称为**共同价值**（common value），它假设所有投标者对物品都有一个估值，但当他们出价时，他们不知道物品的真正价值。每个投标者在参与拍卖之前，对物品价值拥有一个私人观测信号，这就是在出价时他所知道的。那么每个投标者的类型就是对这个共同价值拥有的私人观测信号。一批艺术家经销商竞拍一些艺术品，再倒手卖掉赚钱就是一个例子。真正的价值是艺术品在未来的卖价。

在此我们只是象征性地介绍了两种信息集。在本章附录里，我们提供了一个更完美的均衡——出价行为模型（要利用到微积分知识）。

◯ 案例研究

独立私人价值和隐藏你的出价

1994 年 11 月 11 日，比尔•盖茨在拍卖会上竞拍到了《莱斯特法典》（*the Codex Leicester*）这部科技历史文集。这个命名奇特的物品是将近 500 年前达芬奇所著的关于天文学的一本书。从 1717 年开始伯爵莱斯特和他的后裔拥有这部作品的所有权，直到 1980 年，实业家阿曼德•哈默（Armand Hammer）买下这部作品。哈默逝世后，他的财产全被拍卖。比尔•盖茨以 3 080 万美元的价钱创下非文学作品的拍卖纪录。

让我们假设一件物品，比如《莱斯特法典》正在被拍卖且有两个投标者：比尔和夏琳。拍卖商决定以一级密封价格拍卖的形式进行拍卖，这就要求两个投标者同时提交自己的出价，出价最高的获得这部作品并以自己的出价付钱给拍卖商。我们假设如果两个投标者出价相同，那么拍卖商就通过抛硬币的方式决定谁是赢标者。

每个参与者的类型就是在拍卖会上对物品的估价，为了使得拍卖变得简单，我们假设有两种可能的类型：出价 50，概率为 0.6；出价 100，概率为 0.4。“自然”一开始选择每个参与者的类型。假设他们的类型是独立的，比尔和夏琳都出低价 50 的概率就为 36%（因为 0.36=0.6×0.6），两者都出高价 100 的概率为 16%，一个出高价 100 另一个出价 50 的概率为 48%。每个投标者都只知道自己的估价，然后两个投标者同时提交自己的出价。我们假设竞拍价钱每次增量以 10 为单位：10，20，30，… 如果他是赢标者，该投标者的收益就是他的估价减去他的出价（在一级密封价格拍卖中，出价就等于物品的价格），而如果没中标，他的收益就为 0。

先分析拍卖中一个投标者估价 50，另一个投标者估价 100 时的情况。假设两个投标者都采取以下策略：若估价 50，就出价 40；若估价 100，就出价 60。让我们分析这个对称策略组合是否是一个贝叶斯纳什均衡。为了得到分析结果，我们推导比尔关于夏琳出价的信念。假设比尔认为夏琳会采取上述策略，且他相信夏琳估价为 50 的概率为 60%，这样夏琳的出价为 40；夏琳出价 60 的概率就为 40%。有了这些关于夏琳出价的信念，我们就可以测评出比尔出价的最优策略。

假设比尔估价为 50。他知道夏琳或者出价 40，或者出价 60。如果他的策略为出价 40，那么他得到的期望收益就是：

$$0.6\times0.5\times(50-40)+0.4\times0=3$$

这个表达式是这样推导出来的：夏琳估价为 50 时，出价 40 的概率为 0.6。因为比尔出价和夏琳相同，那么他就有 50% 的可能赢得拍卖品。（记得当两者出价相同时，我们假设拍卖商通过抛硬币的方法来决定赢标者。）这样他得到的收益为 10（= 50 – 40），因为拍卖品价值 50 而他只付出了 40 的价格。

于是，要使出价40成为最佳策略，那么另一个投标者得到的期望收益必须小于或等于3。如果投标者出价低于40，那么他肯定不会中标，这样他的收益为0。如果他出价50，那么他的收益还为0，因为他既没输也没赢，出价等于他的估价。如果他出价高于50，那么他得到的收益或者为0（如果他输了），或者为负值（如果他赢了）。所以，如果比尔的估价为50，那么出价40是最佳策略。

现在假设比尔的估价为100。这样他的出价就为60，得到的期望收益为：

$$0.6\times(100-60)+0.4\times0.5\times(100-60)=32$$

要使比尔出价60为最优策略，那么其他的投标者得到的期望收益不得高于32。让我们证明我们是如何很快剔除那些出价低于40和出价高于70时的策略。如果比尔出价低于40，那么他肯定不会中标（因为夏琳或者出价40或者出价60），他得到的收益就为0，这个策略明显劣于出价60得到收益32的策略。如果比尔出价70，那么他肯定会赢标，因为夏琳出价不会高于60。所以，他的收益就为30。出价高于70只会降低他的收益，因为赢标的可能性不会增加了，出价70时赢的概率已经为1，出价越高仅仅意味着付越高的价钱给拍卖商。

上述分析为我们阐释了相比出价60来说，出价40，50和70也是可能的选择。得到的收益分别如下：

出价40得到的收益 $=0.6\times0.5\times(100-40)+0.4\times0=18$

出价50得到的收益 $=0.6\times(100-50)+0.4\times0=30$

出价70得到的收益 $=0.6\times(100-70)+0.4\times(100-70)=30$

注意这些期望收益都低于出价60时得到的期望收益32。于是，如果投标者1的估价为100，那么出价60就是最优策略。同样，这个分析适用于夏琳。所以，当投标者的估价为50时出价40，以及当投标者估价为100时出价60就是一个贝叶斯纳什均衡。

这就出现了一个很有意思的现象，即当投标者估价高时，提交的出价为60，大大低于估价100。出价比你的估价低就叫做"隐藏你的出价"。若你出价最高且赢标的话，出价越低，你得到的收益越高。然而，更低的出价也可能降低赢标的概率。这两个因素之间的权衡决定了投标者的最佳出价策略。

表10—1描述了当投标者估价为100时的权衡情况。出价为40时，这个投标者有30%的概率获胜，因为另一个投标者估价50出价40的概率为60%，如果另一个投标者也出价40，那么拍卖商就抛硬币来决定谁是赢标者。收益为60（= 100 − 40）时，这个收益值就相当高。相反，如果出价更高的话，比如50，那么这样收益就降到50，但是赢标的概率增加至60%（因为如果另一个投标者估价低的话，他肯定赢标）。于是我们得出这样的结论：出价50得到的期望收益30高于出价40得到的期望收益18。最优出价策略就是出价60，我们注意到这既不是赢标概率最高也不是收益最高的情况。但是，它权衡这两个因素使得投标者得到的期望收益最高。

表10—1　隐藏你的出价，高估价的投标者

出价	赢标的概率	投标者赢标的收益	期望收益
30	0	70	0
40	0.3	60	18
50	0.6	50	30
60	0.8	40	32
70	1	30	30
80	1	20	20

共同价值和赢者的诅咒

想象你自己是一家石油公司，比如说美国德士古公司（Texaco）的高级经理，你正在考虑购买石油的钻探权，以在墨西哥湾的某些地区开采石油。为了得到这个地区的石油总量预计值，你决定雇用一家工程公司来估计石油钻探权的真值。有两家公司可选：A公司和G公司。据行业小道消息，你已经了解到其中一家倾向于积极估价，另一家倾向于消极估价，但你不知道到底哪家是倾向积极估价的，所以你认为每家公司积极估价的概率为$\frac{1}{2}$。所以，G公司高估这个地区的石油总量和A公司低估这个地区的石油总量的概率为$\frac{1}{2}$；G公司低估这个地区的石油总量和A公司高估这个地区的石油总量的概率也为$\frac{1}{2}$。

令v表示石油钻探权的真值，假设v分布在{10, 11, ..., 100}之间。那么积极估价的公司提供的钻探权的估价（或者为观测信号）为：$s=v+2$；而消极估价的公司提供的钻探权的估价为：$s=v-2$。假设你选择G

公司，得到的估价为 s。因为你不知道 G 公司是积极估价还是消极估价，你所知道的是钻探权的真值是 $s-2$（当 G 公司是积极评价时）的概率为 50%，钻探权的真值是 $s+2$（当 G 公司是消极评价时）的概率为 50%。比如，如果观测信号是 10，那么钻探权的真值为 12 的概率为 $\frac{1}{2}$，为 8 的概率也为 $\frac{1}{2}$。更具体地来说，在收到观测信号 s 后，你对钻探权的真值的信念为：

$$\text{Prob}(v|s)=\begin{cases}\frac{1}{2} & v=s-2\\ \frac{1}{2} & v=s+2\end{cases}$$

Prob $(v|s)$ 是你在收到观测信号 s 后赋予钻探权 v 真值的概率。

假设德士古公司是唯一一家与石油钻探权拥有者谈判的公司。拥有者定价且提出买卖自由的策略。如果德士古公司很看重自己期望收益达到最大值，那么他愿意出价的最大值是多少呢？钻探权的期望值为：

$$\left(\frac{1}{2}\right)(s-2)+\left(\frac{1}{2}\right)(s+2)=s$$

所以如果德士古公司出价为 p，得到的期望收益就是 $s-p$。那么德士古公司愿意出的最高价就是 s。尽管这个观测信号被证明既不是太高也不是太低，但它确实反映了钻探权的真值，所以德士古公司愿意以观测信号为依据出价。如果买价 $p<s$，那么德士古公司的期望收益为正值。

现在假设埃克森美孚公司（Exxon）也与石油钻探权拥有者谈判。既然有两家感兴趣的公司，拥有者决定实施一级密封价格拍卖。德士古公司该如何出价呢？让我们来证明他出价等于 s 或略低于 s 时都会赔钱！

在模拟这个涉及两个投标者的贝叶斯博弈时，我们假设“自然”先行选择石油钻探权的真值，这个真值分布在 {10, 11, …, 100} 之间。给定钻探权的真值 v，每个投标者的类型是他自己的观测信号。假设埃克森美孚公司雇用其中一家工程公司，而德士古公司雇用另一家公司。那么，德士古公司得到的观测信号为 $v+2$ 的概率为 $\frac{1}{2}$（如果最终雇用的是积极评价公司），埃克森美孚公司得到的观测信号为 $v-2$ 的概率就为 $\frac{1}{2}$；德士古公司得到的观测信号为 $v-2$ 的概率为 $\frac{1}{2}$（如果最终雇用的是消极评价公司），埃克森美孚公司得到的观测信号为 $v+2$ 的概率就为 $\frac{1}{2}$。所有的这些信息都是共同知识。根据一级密封拍卖的规则，每个投标者在了解自己类型（也就是他自己的观测信号）之后，从集合 {1, 2, …, 100} 中提交出价。出价最高者获得物品的拥有权且支付和他出价等同的价钱给拍卖商。如果两家公司出价相同，那么每个公司获胜的概率为 50%。若没赢标，得到的收益为 0，如果赢标，得到的收益为 $v-b$，这里的 b 是投标者的出价。所以，投标者只有在拍卖结束后才能知道钻探权的真值（或更准确地来说，在开采石油后才能知道钻探权的真值）。

这里我们的目的不是为了证明均衡出价是什么样的，而是为了证明它不是什么样的。思考出价为 $s-1$ 这个对称策略，这样投标者出价只比钻探权的真值少一点。注意这样的出价规则有两个合理的特征：（1）投标者出价比钻探权的期望值低（我们已经证明为 s）；（2）当投标者的观测信号越高时，他出价越高。

假设埃克森美孚公司遵循同样的规则，那么德士古公司从这个规则中得到的期望收益是多少呢？为了使这个出价成为一个贝叶斯纳什均衡，对每个类型的投标者来说，$s-1$ 这个出价都必须使他们的期望收益最大化。比如，当 $s=20$ 时，德士古公司的出价为 19。在计算期望收益时，我们使用这个信念，即钻探权真值为 22 的概率为 50%，为 18 的概率为 50%。如果钻探权的真值为 18，那么德士古公司就有一个高估值，埃克森美孚公司就有一个低估值（具体来说，为 16），这就意味着德士古公司出价 19 会淘汰埃克森美孚公司出价 15，这样德士古公司就赢得钻探权。现在有一个坏消息：德士古公司付出的代价太高。他得到的收益为−1，因为钻探权的真值为 18 而他出价 19。现在假设真值为 22 而非 18。因为德士古公司的观测信号为 20，埃克森美孚公司的观测信号为 24，所以出价 23，这个出价高于德士古公司出价的 19。所以德士古公司不会中标。

总结这两种情况，我们知道（当他的观测信号为 20 时），德士古公司出价 19 得到的期望收益为：

$$\left(\frac{1}{2}\right)\times(18-19)+\left(\frac{1}{2}\right)\times 0=-\frac{1}{2}$$

更一般地来说，当两个投标者遵循相同的出价规则时，德士古公司得到的期望收益为：

$$\left(\frac{1}{2}\right)\times[(s-2)-(s-1)]=-\frac{1}{2}$$

所以，尽管一个公司出价比他的观测信号低且这个信号值和钻探权的真值相同，这个公司还是会赔钱！

当德士古公司是唯一一个潜在的买主时，出价低于自己的观测信号就意味得到的期望收益为正值，因为他的出价比钻探权的期望值低。在拍卖结束的时候，他最

终得到的收益可能为正值（如果估价低），可能为负值（如果估价高），但德士古公司肯定希望得到正的收益值。相反，如果德士古公司在拍卖中出价低于自己的观测信号，正如我们前面证明的那样，他得到的收益可能为0（如果他没有赢标），也可能为负值（如果他赢标）。这里起关键作用的是在拍卖界存在的最令人畏惧的力量：**赢者的诅咒**（winner's curse）。

赢者的诅咒指的是这样一种现象，即赢标不一定是好消息。当物品的价值对所有的投标者来说是共同知识，且每个投标者得到的都是不准确的估价时，就可能发生赢者的诅咒现象。如果所有的投标者遵循相同的出价规则，就会提高投标者的估价，那么这个事实就意味着赢标者的估价是所有投标者中估价最高的。所以，当不知情的投标者出价高时，赢标的事实就反映了他对这个物品的真正价值有最乐观的估计，而这个估计值，通常情况下会高于物品的真正价值。（在当前的这个例子中，赢标就意味着你雇用的工程公司给你透露的观测信号一定大于物品的真值。）赢得的拍卖品透露给你的信息就是你的估值大大高于物品的真值，所以就隐藏着坏消息。

令人欣慰的是，博弈论为赢者的诅咒提供了一剂良药。技巧在于在这样的假设下出价：你的观测信号是所有的投标者中最高的，这样你才可能赢得拍卖。所以，如果证明你的观测信号是所有投标者中最高的，你就会确定物品的期望价值，然后，你会结合这个期望价值合理地出价。在前面的例子中，这个策略提示投标者出价不应高于$s-2$。因为如果证明你有更高的观测信号，你雇用的肯定是一个积极评估的公司，这样物品的真值为$s-2$。高于这个真值的出价都会导致亏损。准确地说，你应该如何出价是一个很棘手的问题，这将在附录2里进一步解释，附录2描述了研究共同价值拍卖的另一种方法。

实际上，投标者会从赢者的诅咒中蒙受损失吗？一项最近的研究调查了大型石油公司如埃克森美孚、英国石油和德士古公司在这类情况下精明的出价的案例。即将拍卖的是美国政府持有的距离海岸3~200英里的矿产权。拍卖的是这5 000英亩的土地，公司竞拍是为了获得石油的开采权。证据表明大部分公司的管理者确实是理性的决策者，因为大多数情况下他们都避免了赢者的诅咒现象的发生。如果他们认为自己对物品的真值估计是所有投标者中最高的话，他们的出价大约是石油钻探权期望收益的三分之一。这就是说，在提交最高竞价赢得拍卖时，这个出价就会产生一个正的期望收益。一个值得关注的例外是德士古公司，他的出价不仅使他遭受赢者的诅咒，且他的出价比石油钻探权的估计值更高！在这个例子中，德士古公司出价高于自己的观测信号s，这个策略肯定会导致亏损！真希望德士古公司或者换掉那些管理者，或者让他们补修博弈论这堂课。

10.4 委员会和陪审团的投票

另一个涉及私人信息的博弈是一群人必须达成某种共识。开会前每个人都拥有私人信息，达成共识的手段就是投票。首先，我们思考当人们怀有不同的信息并同时提交选票的策略问题。接着，我们来探索在一个法庭现场，陪审团对被告的罪行持有不同的观点，他们通过序贯投票来决定最后的裁决。

10.4.1 弃权投票策略

都说投票是公民应尽的义务。但通常情况下，投票就会产生最好的结果吗？实际上，在选举日，你在咖啡厅悠闲地喝着咖啡，放弃你的投票权利，这对社会是否最好呢？正如我们即将证明的，在某些情况下，你自己的兴趣或是社会的某种力量都会使你在政治问题上保持沉默。

思考一个三人组成的委员会必须决定是继续实施现阶段政策还是实施新政策的博弈。如果这三名委员对新政策有足够的了解，且他们对新政策有同样的偏好，那么他们就会一致支持或一致反对新政策。问题在于他们中的有些委员对新政策的了解不够全面。如果委员会选择的政策被认为是最佳的，每个委员得到的收益为1，而如果选择了错的政策，每个委员得到的收益为0。

每个委员有两种类型：了解新政策的和不了解新政策的。对新政策不够了解的委员就不能确定哪个政策更好。更具体地来说，我们假设现阶段政策更好的概率为p，这里$\frac{1}{2}<p<1$。所以，不需要任何附加消息，不了解新政策的委员一定会选择现阶段的政策，因为这样得到的期望收益为$p\times1+(1-p)\times0$，即p，而选择新政策得到的期望收益为$p\times0+(1-p)\times1$，即$1-p$。相反，如果这个委员对新政策有足够的了解，那么她就知道哪个政策更好。

“自然”先行决定哪个政策更好，以及每个委员所了解新政策的情况。为了保证对新政策不了解的委员和“自然”的意向一致，假设“自然”选择现阶段政策更好的概率为 p。在决定了哪个政策更好之后，新政策被委员了解的概率为 q。于是，一个委员了解新政策类型的概率就为 q。他对新政策不了解的概率就是 $1-q$，这样她就不确定哪个政策更好，但更倾向于认为现阶段政策更好（如 q 的值所表示）。

在“自然”决定这四项之后（决定哪个是更好的策略以及三个委员的类型），这三个委员同时投票。每个委员可以支持现阶段的策略，可以反对现阶段的策略，也可以弃权。如果大多数委员反对现阶段的政策，那么实施新政策；如果大多数委员支持现阶段的政策，那么继续实施现阶段的政策。在策略实施之后，委员就会了解到他们做了正确的选择（得到的收益为 1），还是做了错误的选择（得到的收益为 0）。

在这个贝叶斯博弈中，一种策略就是一组行动：当一个委员是了解新政策的类型时，是否去投票，如何投票，以及当一个委员是不了解新政策的类型时，是否去投票，如何投票。一个对称贝叶斯纳什均衡是这样一种策略：给定其他两个委员采取相同的策略，以及对于其他成员拥有哪个政策更好的知识的信念，不论这个参与者是什么类型，采取这个策略都使她得到的收益最高。

让我们先讨论当一个委员的类型是了解新政策的时候，她投票选择自己认为是最好政策的策略就是弱占优策略的情况。通过投票选择自己认为最好的政策，委员 1 者不能影响投票结果（因为委员 2 和委员 3 的投票和委员 1 的投票相同），或者是她认为最好的政策被实施（当委员 2 和委员 3 在新旧政策上有分歧的时候）。在第一种情况下，通过投票选择自己认为最好的政策，委员 1 得到的收益不会更低（相比投票自己认为不好的政策和弃权来说），在第二种情况下，她得到的收益严格更高。我们总结出：对一个了解新政策的类型来说，投票选择自己认为更好政策的策略是一个弱占优策略。

现在我们开始研究更有趣的情况。假设一个委员不知道哪个政策更好。投票是最优策略，还是弃权是最优策略？

让我们首先思考一个对称策略组合，当委员是了解新政策类型时，投票选择自己认为最好的政策；当委员是不了解新政策类型时，投票现阶段的政策。（记得不了解新政策类型的成员认为现阶段政策更好的可能性更大。）这就意味着委员 1 总是希望委员 2 和委员 3 投票（不管他们是不是了解新政策的类型）。只有当委员 2 和委员 3 对新旧政策有分歧时（一个选择现阶段政策，另一个选择新政策），委员 1 的投票才起决定意义。然而，这也意味着有些人会投票选择不是最好的政策，这就使得这个参与者必须被“告知”哪个政策才是最好的。（记得我们刚刚讨论过在对称策略组合中，不了解新政策的类型通常会选择投票现阶段的政策。）所以如果委员 1 是不了解新政策的类型，她会投票现阶段的政策，那么社会实施了错误的政策。相反，如果她选择投票新政策，结果会比这个好。所以对不了解新政策的选民来说，投票现阶段的政策不是最优策略。也就是说，如果委员 2 和委员 3 是不了解新政策的类型且投票选择现阶段的政策，那么如果委员 1 也是不了解新政策的类型，她投票选择现阶段的政策就不是最优策略。这个推理证明了当参与者是不了解新政策的类型时，投票选择现阶段政策不是一个对称贝叶斯纳什均衡，尽管她认为现阶段的政策更好的可能性更大。

另一个可能成为贝叶斯纳什均衡的对称策略组合是：当参与者是不了解新政策类型时，投票反对现阶段的政策；当参与者是了解新政策类型时，投票选择最好的政策。然而，我们同样可以证明这不是一个贝叶斯纳什均衡。证明过程和前面的相同。

最后一个可能成为贝叶斯纳什均衡的对称策略组合是：当参与者是不了解新政策的类型时，选择弃权；当了解了新政策之后，投票选择最好的政策。假设委员 1 采取这个策略，同时她期望委员 2 和委员 3 采取同样的策略。如果委员 2 或（和）委员 3 投票，那么根据她们的策略，她们一定被“告知”了新政策的情况。这种情况下，最好的政策就会被实施，这就使得委员 1 通过弃权得到最大收益 1。如果委员 2 和委员 3 都不投票，那么她们都没被“告知”新政策。于是通过弃权，委员 1 使得社会又继续实施现阶段的政策。如果委员 1 去投票，对她来说最优策略就是投票选择现阶段的政策，因为她认为现阶段政策更好的可能性要大。这样，现阶段的政策就导致委员 1 或者弃权或者投票现阶段政策。所以，当参与者不了解新政策的类型时，最好的策略就是选择弃权，给定其他两个委员在不了解新政策时也弃权。

总之，对一个委员会成员来说，当不了解新政策时

弃权和当了解新政策时投票选择最好的政策就是一个对称贝叶斯纳什均衡。尽管投票不需要成本，有时候弃权却是最好的策略。这个命题的推理过程很简单：如果这个参与者不了解新政策，那么就弃权让了解新政策的参与者决定最后的结果。所以，只有不了解情况的人才有资格弃权，这就意味着如果你在选举日当天去咖啡馆闲逛，你应该读一读《人物》杂志而不是看《纽约时报》。

10.4.2 陪审团室的序贯投票机制

现在假设你是一个陪审员。法庭询问已经结束，你和其他陪审员正要离开陪审团室。陪审团团长以投票的方式开始审议。所有的陪审员围桌而坐，轮流宣布被告有罪或无罪。让我们以投票顺序来为陪审员编号，所以陪审员 1 第一个宣布她的表决，接着是陪审员 2 宣布，依此类推直到陪审员 12 宣布她的表决。

基于审判过程中的辩论，每个陪审员在进入陪审团室之前对被告是有罪还是无罪已经有一个大体的评判。这个评判——她们意识到可能不是准确的观测信号，就是陪审员的类型。[①]现在假设在投票时，陪审员希望在她了解的所有情况基础上进行评判。对陪审员 1 来说，她所了解的所有情况就是在审判时所做的评判。相反，在目睹了陪审员 1 投票之后，陪审员 2 就可以推断出陪审员 1 的评判。如果情况是这样的话，那么陪审员 2 不仅拥有自己的评判结果，也知晓陪审员 1 的评判结果。类似的，陪审员 3 可以推断陪审员 1 和陪审员 2 的评判结果，所以陪审员 3 拥有三份评判结果。假设一个陪审员更倾向于（1）如果评判有罪的人数多于评判无罪的人数，就投票有罪；（2）如果评判无罪的人数多于评判有罪的人数，就投票无罪；（3）如果评判有罪的人数和评判无罪的人数相同，那么根据自己的评判结果投票。在情况（3）中，一个陪审员的评判结果在评估过程中占了更大的分量。[②]

对陪审员 1 来说，贝叶斯策略赋予她投票的策略，有罪或无罪，取决于她的评判结果（或类型）：有罪或无罪。对陪审员 2 来说，贝叶斯策略赋予她投票的策略取决于他自己的评判结果和陪审员 1 的投票结果。对陪审员 3 来说，贝叶斯赋予他投票的策略取决于他自己的评判结果，以及陪审员 1 和陪审员 2 的投票结果，依此类推。现在让我们构建贝叶斯纳什均衡。

我们很容易求解陪审员 1 的最优策略，因为只需要考虑她自己的评判结果。如果在审判结束后她认为被告有罪，那么她就投票有罪；如果她认为被告是清白的，她就投票被告无罪。

陪审员 2 观察到陪审员 1 的投票结果，依此推导陪审员 1 的策略。尽管他无法观测到陪审员 1 的评判结果——是陪审员 1 的私人信息，但他能够推导出来，因为陪审员 1 是根据自己的评判结果来投票的。所以，如果陪审员 1 投票有罪，那么陪审员 2 可能推断出陪审员 1 的评判结果是被告是有罪的。现在假设陪审员 2 自己的评判结果是被告有罪。那么陪审员 2 毫无疑问投票被告有罪，因为他有两条被告有罪的评判结果。相反，如果陪审员 2 自己的评判结果是被告无罪，那么他有一条有罪评判结果（陪审员 1 的评判结果）和一条无罪评判结果（他自己的）。如前面陈述的，在这种情况下陪审员倾向依据自己的评判结果来投票，所以陪审员 2 投票被告无罪。不论陪审员 2 的评判结果和推断的陪审员 1 的评判结果是否一致，我们得出陪审员 2 的贝叶斯纳什均衡策略是根据自己的评判结果来投票。陪审员 1 如何投票并不影响陪审员的投票结果。

至此，发生的一切都平淡无奇：当且仅当前两个陪审员自己的评判结果和证据都表明被告是有罪的，他们才会投票被告有罪。那我们还有必要观察陪审员 3 的投票行为吗？确实有必要。陪审员 3 观测到陪审员 1 和陪审员 2 的投票，根据我们刚刚推导的均衡策略，每个陪审员的投票结果都和自己的评判结果一致。有两种情况需要考虑：（1）陪审员 1 与陪审员 2 的投票结果相同；（2）陪审员 1 与陪审员的投票结果不同。让我们逐一对这两种情况进行分析：

■ 假设陪审员 1 和陪审员 2 的投票结果不同。那么陪审员 3 就应该依据自己的评判结果进

① 将这个博弈恰当地定义为贝叶斯博弈，需要明确出陪审员判定有罪类型的概率（即他进入陪审员室持有罪判定的信念）或者判定无罪类型的概率。结果证明这些概率是多少并不重要，因此可以不必一一列举。——作者注

② 尽管在这里我们还没有明确提到收益值，但我们很容易就能介绍收益值的情况。比如，假定被告有罪的评判结果条数多于被告无罪的评判结果条数，若她的投票结果为“被告有罪”，得到的收益为10，若她的投票结果为“被告无罪”，得到的收益为5。于是，在这种情况下，投票被告有罪使她得到的收益最大。——作者注

行投票。比如，假设陪审员 1 投票有罪，陪审员 2 投票无罪。如果陪审员 3 自己的评判结果是被告有罪，那么就有两条被告有罪的评判结果，那么陪审员 3 就应该投票被告有罪。如果陪审员 3 的评判结果是被告无罪，那么陪审员 3 就应该投票被告无罪。所以，当陪审员 1 和陪审员 2 的投票结果有分歧时，陪审员 3 应该依据自己的评判结果来投票。

- 现在假设陪审员 1 和陪审员 2 的投票结果相同。如果他们都投票被告有罪，那么根据陪审员 3 对这两个陪审员策略的推断，她推导出陪审员 1 和陪审员 2 审判结束后的评判结果都是被告有罪。如果陪审员 3 自己的评判结果是被告有罪，显然，她应该投票被告有罪。但即使是陪审员 3 自己的评判结果是被告无罪，她还是投票被告有罪，因为有两条被告有罪的评判结果，只有一条被告无罪的评判结果。更具体地来说，如果陪审员 1 和陪审员 2 的投票结果相同，陪审员 3 投票应该和陪审员 1 与陪审员 2 的投票相同，因为这两条相同的评判结果会使陪审员 3 的投票结果失效。在投票时，陪审员 3 就应该忽略自己的评判结果。

那么，当陪审员 1 和陪审员 2 的投票结果相同时，陪审员 3 的贝叶斯纳什均衡策略就使她依据陪审员 1 和陪审员 2 的评判结果来投票。只有当陪审员 1 和陪审员 2 的投票结果不同时，陪审员 3 投票的最优策略才是依据她自己的评判结果进行投票。

用这种方法，我们可以构建陪审员 4 到陪审员 12 的贝叶斯纳什均衡投票策略。但我们直接推导那些例外的情况。假设审判结束，当 12 个陪审员走进陪审团室时，10 个陪审员都认为被告无罪，只有两个陪审员相信被告被指控的罪名是成立的。那么，尽管拥有被告无罪的先验经验，最开始的投票结果使得所有的陪审员都投票被告有罪！尽管每个陪审员都依据信息理性地进行投票！

证明过程如下：假设这两个认为被告有罪的刚好是陪审员 1 和陪审员 2，他们两个最先投票。利用我们刚刚推导出的均衡投票策略：显然，陪审员 1 投票被告有罪，陪审员 2 也投票被告有罪，因为根据我们刚刚推导的，陪审员 2 的最优策略是和他自己的评判结果一致，不受陪审员 1 投票结果的干扰。当轮到陪审员 3 投票时，她也会投票被告有罪，与她的策略规定一致。从陪审员 1 和陪审员 2 的投票结果来看，她可以推断出陪审员 1 和陪审员 2 都认为被告是有罪的，不论陪审员 3 如何投票都不能改变最终的投票结果。所以，她会忽视自己的评判结果，投票被告有罪，因为根据陪审员 3 的策略，如果陪审员 1 和陪审员 2 投票结果一致，她就投相同的票。当轮到陪审员 4 投票时，他知道陪审员 3 忽视自己的评判结果，因为当陪审员 1 和陪审员 2 的投票相同时，陪审员 3 的策略使她的投票和陪审员 1、陪审员 2 的投票结果一致。所以，陪审员 4 所能推断的只是陪审员 1 和陪审员 2 的评判结果：都认为被告有罪。陪审员 4 和陪审员 3 的情形一样，于是也投票被告有罪。这个理论同样适用于陪审员 5，6，…，12。所以，陪审员 12 也投票被告有罪！然而，这 12 个陪审员不全是依照自己的评判结果投票，因为陪审员 3，4，5，…，12 在投票时都忽视了自己的评判结果。裁决被告有罪仅仅因为前两个陪审员认为被告是有罪的。

寓审员 3，4，5，…，12 就是羊群行为（sheeple）[①] 的代表吗？相反，他们并不是不经大脑思考简单模仿前两个陪审员的投票行为，而是对他们前面陪审员的投票结果进行了复杂的推理。那 10 个陪审员依据拥有的最多信息来投票。这就需要他们推导前面投票者的评判结果。

这种行为被称为“信息瀑布”或“羊群效应”，它指在某种情况下，人们的行为有先后顺序，前面几个参与者行为透露的信息主导了群体中每个参与者的观测信号，使得其他参与者无视自己的观测信号；即其他参与者的行为仅仅依靠前面几个参与者的观测信号，他们做出的决定完全和自己的私人信息无关。一旦一个参与者无视自己的私人信息，他的行为提供的信息就没有任何

① 这个单词是由“sheep”和“people”组合而来的。它是指有些人只是简单地不加思考地尊循别人的做法，尽管有时候他认为那样做是不恰当的。——作者注

额外价值，这同样适用于在他后面行动的参与者。所以我们发现高度一致的行为，比如所有的陪审员投票被告有罪，仅仅取决于一小部分参与者的观测信号。在这个例子中，这个行为依赖的只是前面两个陪审员的观测信号，他们依据证词和证据判断被告有罪。

这里，我们得出一个重要的结论：投票程序至关重要。比如，我们假设不是序贯投票，而是陪审团团长规定每个陪审员在一张纸上写出自己的投票结果并同时投票。每个陪审员的最优策略就是审判结束后根据自己的评判结果写出投票结果。那么，因为我们已经假设 10 个陪审员认为被告无罪，投票结果就为无罪：有罪 =10：2，这个结果和之前的有罪裁决结果完全不同！

本章总结与练习题

本章总结

本章讨论了多个策略集的常见的和重要的特征：一个参与者可能知道他或她自己知道而其他参与者不知道的信息。一个参与者的收益为私人信息时，这种情况就会发生。如我们刚开始接触到的，这个博弈不是共同知识，因为一个参与者不知道其他参与者的收益。我们称这类博弈为**不完全信息博弈**。求解不完全信息博弈的关键和技巧就在于把不完全信息博弈转换为不完美信息博弈。博弈一开始是由随机力量决定的，我们称之为“**自然**”。“自然”决定每个参与者的**类型**，这里类型包含一个参与者拥有的私人信息。这种“自然”参与的博弈也称为**贝叶斯博弈**，在这类博弈中，因为博弈一开始的时候，所有参与者都不知道自己的类型，也就没有私人信息。所有参与者知道的是“自然”将会决定每个参与者的类型以及“自然”赋予每个参与者类型空间的概率。

本章求解贝叶斯博弈利用的概念是贝叶斯纳什均衡。贝叶斯纳什均衡类似于纳什均衡，它是指在给定自己的类型以及其他参与者的类型与战略选择之间关系的条件下，使得自己的期望收益最大化的一种策略。但这个定义还需要补充两个方面。第一，一个参与者不知道其他参与者的类型，所以不知道其他参与者将采取什么行动。然而，如果一个参与者（准确地）推测另一个参与者的策略和与这个参与者类型相关的信念，他就可以构建这个参与者将如何采取行动的信念。第二，在给定这些信念后，一个参与者的策略必须为参与者每一种可能的类型规定一个最优行动方案。

在本章里，我们借助几个例子探讨了均衡状态下参与者的行为。这些例子包括：枪战博弈，在这个博弈中，一个熟练的枪手并不知道对手的底细；委员会投票博弈，其中，每个参与者都不知道其他参与者的最终目的，委员会的每个成员也不知道其他委员的信息拥有量；拍卖博弈，在这个博弈中，只有拍卖商知道物品的真实价值。这只是拥有私人信息博弈策略的几个例子。每天你都会碰到很多这类情况。

值得你关注的是最后一个例子，它涉及的是陪审团室的序贯投票机制。在这章的所有其他例子中，拥有私人信息的参与者或者和其他参与者同时采取行动（拍卖博弈），或者在其他参与者采取行动之后采取行动（慕尼黑协定博弈），这就意味着其他参与者没有机会从某个参与者的行为推测其拥有的私人信息。但在陪审团博弈中，一个观测到其他陪审员投票行为的陪审员在投票时，可能受到他所推断其他陪审员评判结果的影响。这正是我们在下一章即将讨论的信息集。你将会学习到不同信息集之间的微妙差别。

PDG

练习题

1. 格雷格正在考虑是否邀请马西娅。然而，格雷格不确定马西娅是否喜欢他，且如果是被拒绝的结果，他宁愿不约马西娅。马西娅喜不喜欢格雷格只有她自己知道，即这是马西娅拥有的私人信息。所以，她对格雷格的喜欢与不喜欢构成了她的类型。格雷格并不拥有任何私人信息。假设马西娅喜欢格雷格的概率为 25%。这个贝叶斯博弈如图 10—6 所示。格雷格应不应该邀请马西娅与他约会呢？求出这个贝叶斯纳什均衡来回答这个问题。

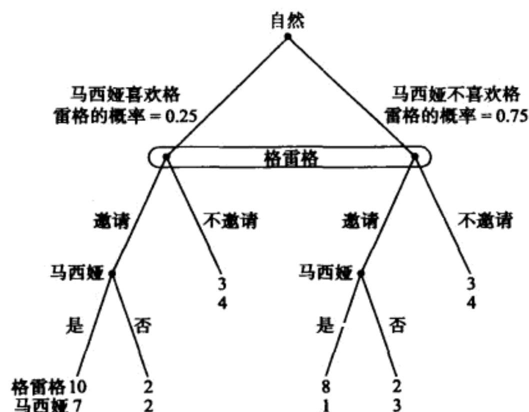


图 10—6

2. 思考 B 和 C 之间的枪战博弈。两个参与者都拥有自己使用六发手枪能力的私人信息。“自然”先行决定每个枪手的能力。他可以是快发枪，也可以是慢发枪。B 为快发枪类型的概率为 50%，C 为快发枪类型的概率为 60%。在每个枪手了解自己的类型后——他选择开枪或僵持。如果都选择僵持，那么收益为 50。如果都开枪，（1）他们的类型相同（或都为快发枪，或都为慢开枪），那么每个枪手得到的收益为 20；（2）如果他们是不同的类型，那么快发枪的枪手得到的收益为 30，慢发枪的枪手得到的收益为-40。如果一个开枪，另一个僵持，（1）如果他们的类型相同，那么发枪的枪手得到的收益为 30，僵持的枪手得到的收益为-40；（2）如果开枪的是快开枪，另一个是慢发枪，那么快开枪的枪手得到的收益为 30，慢发枪的枪手得到的收益为-40；（3）如果开枪的是慢开枪，另一个是快发枪，那么每个参与者得到的收益为 20。如果至少有一个选择开枪，就会发生枪战。

a. 存在一个贝叶斯纳什均衡使枪战一定会发生吗？（就是说，不论参与者的类型是什么，他们都选择开枪。）

b. 存在一个贝叶斯纳什均衡使枪战一定不会发生吗？（就是说，不论参与者的类型是什么，他们都选择僵持。）

c．存在一个贝叶斯纳什均衡使得只有当参与者是慢发枪时，他才开枪吗？

3. 思考在一个一级密封价格拍卖博弈中，投标者的估价可能是 5，7 和 10，每个发生的概率分别为 0.2,0.5 和 0.3。只有两个投标者，他们的类型由“自然”决定。每个投标者了解自己的类型后，他们同时选择一个出价（这个出价一定是整数）。如果一个投标者没中标，那么他的收益为 0，如果赢得拍卖品，那么他得到的收益为他的估价和出价的差值。

a. 确定当投标者的估价为 5、出价为 4 时，估价为 7、出价为 5 时，以及估价为 10、出价为 6 时的情况是不是对称贝叶斯纳什均衡？

b. 确定当投标者的估价为 5、出价为 4 时，估价为 7、出价为 6 时，以及估价为 10、出价为 9 时的情况是不是对称贝叶斯纳什均衡？

4. 思考在一级密封价格拍卖中，假设只有三个可行的出价：一个投标者可以出价 1，2 或者 3。如果没有赢标，投标者得到的收益为 0。赢标者得到的收益等于他的估价减去他付给拍卖商的价钱（根据拍卖规则，付给拍卖商的价钱就等于他的出价）。每个投标者的私人信息是物品在他看来所值的价钱；所以，投标者的类型就是他的估价。假设只有两种估价，分别用 L 和 H 表示，这里，$H>3>L>2$。假设每个投标者估价为 H 的概率为 0.75。于是我们构建一个贝叶斯博弈：首先，“自然”选择两个投标者的类型。其次，每个投标者了解自己的类型，但并不了解另一个投标者的类型。最后，两个投标者同时提交自己的出价。每个投标者的策略是一组行动：估价高时如何出价和估价低时如何出价。

a. 推导 H 和 L 的条件式，使得投标者估价高时出价 3 和投标者估价低时出价 2 成为一个对称贝叶斯纳什均衡。

b. 推导 H 和 L 的条件式，使得投标者估价高时出价 2 和投标者估价低时出价 1 成为一个对称贝叶斯纳什均衡。

c. 推导 H 和 L 的条件式，使得投标者估价高时出价 3 和投标者估价低时出价 1 成为一个对称贝叶斯纳什均衡。

d. 推导 H 和 L 的条件式，使得投标者不论估价高还是估价低时都出价 1 成为一个对称贝叶斯纳什均衡。

5. 思考图 10—7 中的贝叶斯博弈。"自然"选择参与者 1 的类型，这里类型 H 出现的概率为 p，类型 L 出现的概率为 $1-p$。参与者 1 了解他的类型后，或者选择行动 x 或者选择行动 y。同时，参与者 2 选择行动 a 或选择行动 b。

a. 假设 $p=0.75$，求解贝叶斯纳什均衡。

b. 当 p 为任意值时，求解所有的贝叶斯纳什均衡。

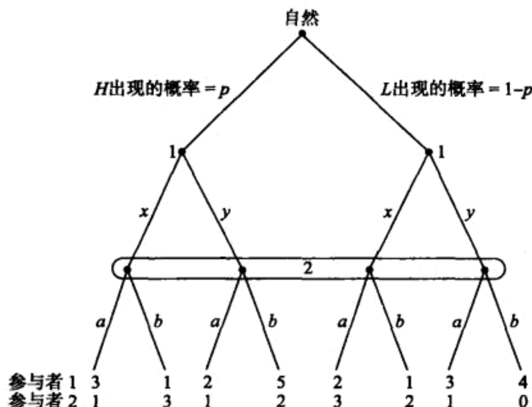


图 10—7

6. 两个美国议员正在考虑参加民主党的美国总统提名竞选。每个候选人进入竞选的个人成本只有自己知道。假设低的个人参加竞选成本 f_L 的概率为 p，高的个人参加竞选成本 f_H 的概率为 $1-p$。所以，类型空间只有两个值。一个候选人的收益取决于他和另一个候选人是否参加竞选。令 v_2 表示当两个候选人都参加竞选时，一个候选人的收益（所以有两个候选人）；v_1 表示当一个候选人参加竞选，另一个候选人没参加竞选时（所以只有一个候选人），参加竞选的候选人的收益；当他没进入竞选时，得到的收益为 0。假设：

$$v_1 > v_2 > 0$$

$$f_H > f_L > 0$$

$$v_2 - f_L > 0 > v_2 - f_H$$

$$v_1 - f_H > 0$$

a. 推导只有当这个候选人参加竞选的个人成本低时，他才参加竞选的策略是对称贝叶斯纳什均衡的条件式。

b. 推导当候选人参加竞选的个人成本低时，一定参加竞选，以及当她参加竞选的个人成本高时，且参加竞选的概率在 0 与 1 之间的策略为对称贝叶斯纳什均衡的条件式。

c. 求解除（a）和（b）以外，其他的贝叶斯纳什均衡。

7. 假设两个国家正准备打仗且同时决定是否进攻。一个国家的军事资源就是这个国家的类型，不同的类型参数影响他们作战取得胜利的可能性。假设类型空间有两个值：p' 和 p''，这里 $0<p'<p''<1$。一个国家为类型 p' 的概率为 q，为类型 p'' 的概率为 $1-q$。思考一个国家类型为 p（或者是类型 p'，或者是类型 p''）的情况。如果一个国家选择率先进攻另一个国家，那么这个国家相信自己可以胜利的概率为 xp，这里 x 取值使得 $p<xp<1$。如果这两个国家都选择进攻，类型为 p 的国家获胜的概率为 p。如果 p 类型的国家选择不进攻，而另外一个国家选择进攻，那么类型为 p 的国家获胜的概率为 yp，这里 yp 取值使得 $0<yp<p$。最后，如果两个国家都选择不进攻，那么就不会发生战争。一个国家的类型越高且如果它比另一个国家先开战，它获胜的概率就越大。当战争没发生时，一个国家得到的收益为 0；获胜时得到的收益为 W，输了战争得到的收益为 L。假设 $W>0>L$。

a. 推导一个国家（不论是什么类型）进攻策略成为一个对称贝叶斯纳什均衡的条件式。

b. 推导只有当这个国家的类型为 p'' 时才进攻的策略成为一个对称贝叶斯纳什均衡的条件式。

8. 思考在一级密封价格拍卖中有三个投标者。当投标者在拍卖会上没有赢得物品时，他得到的收益为 0；当他是赢标者时，得到的收益为他的估价减去他的出价（出价等于他所付给拍卖商的钱）。如果两个或三个投标者都出了最高价，那么随机决定谁是赢标者。假设物品对所有的投标者来说，价值都一样，但是每个投标者有自己的观测信号。"自然"决定物品的真正价值，我们用 v 来表示，v 有三个可能的值：4，5 和 6。每个值被选择的概率为 $\frac{1}{3}$。三个投标者得到的观测信号为 $v-1$，v，和 $v+1$；就是说，一个投标者得到的观测信号太低（$v-1$），另一个投标者得到的观测信号太高（$v+1$），第三个投标者得到的观测信号恰好为 v。每个投标者只知道自己的观测信号，这也是投标者的类型。如果 $v=4$，那么一个投标者得到的观测信号就是 3，另一个投标者得到的观测信号是 4，第三个投标者得到的观测信号是 5。如果

$v=5$，那么一个投标者得到的观测信号是 4，另一个投标者得到的观测信号为 5，最后一个投标者得到的观测信号是 6。如果 $v=6$，那么一个投标者得到的观测信号是 5，另一个投标者得到的观测信号是 6，最后一个投标者得到的观测信号为 7。所以，如果一个投标者得到的观测信号是 5，那么他不知道物品的真正价值是 4（这样他的观测信号最高），5（这样他的观测信号是准确的）还是 6（这样他的观测信号最低）。给定物品的真正价值，每个投标者得到这三个观测信号的机会均等。假设出价的最低增额是 1，那么可行的出价集合是 $\{0, 1, \cdots, 10\}$。

a. 证明表 10—2 所示的对称策略组合是一个贝叶斯纳什均衡：

表 10—2

观测信号	出价
3	4
4	4
5	4
6	4
7	5

b. 证明当投标者的出价严格递增时，不存在对称贝叶斯纳什均衡。也就是说，如果 $b(s)$ 代表一个投标者的出价，给定她的观测信号为 s，且如果：

$$b(7)>b(6)>b(5)>b(4)>b(3)$$

那么这个策略不是一个对称贝叶斯纳什均衡。

9.《新婚游戏》（*the Newlywed Game*）是美国 1966 到 1974 年播放的一档流行直播秀。节目中，四队新人中的一方被问及关于另一方的情况，看他们彼此了解的程度。当妻子在幕后的隔音室时，丈夫回答节目安排的一系列问题，然后妻子回答同样的问题。节目的目的是看丈夫和妻子答案的匹配度。在第二回合，他们角色互换。答案匹配度最高的新人获胜。被问及的问题如：你的岳母（婆婆）经常在哪些事情上提醒你？你的另一半最近有没有说你接吻没有激情？你认为下面哪种最贴切地描述了你新婚之夜的妻子（丈夫），杜松子酒，雪莉酒，还是冰冻果汁鸡尾酒？如果他们的答案一致，那么这对新人得到的收益为 100；如果不一致，那么得到的收益为 0。为了把这个游戏模拟为贝叶斯博弈，假设丈夫的类型空间是杜松子酒，雪莉酒，冰冻果汁鸡尾酒，相应的概率分别为 $\frac{1}{4}$，$\frac{1}{2}$，和 $\frac{1}{4}$。妻子的类型空间也是这三个，相应的概率分别为 $\frac{1}{2}$；$\frac{1}{4}$；$\frac{1}{4}$。参与者的类型互不影响。

a. 求解一个贝叶斯纳什均衡。

b. 求解一个贝叶斯纳什均衡使得丈夫总是说真话。

10. 参与者 1，2 和 3 参加一个合作博弈。每个人的选择策略有三种：A，B 和 C。"自然"决定哪个策略是最优的，每种策略被选中的概率均为 $\frac{1}{3}$。如果所有的参与者都选择的是"自然"决定最优的策略，那么每个参与者得到的收益为 5。如果所有参与者选择同一种策略，但不是"自然"决定最优的那种策略，那么每个参与者得到的收益为 1。如果参与者最终没有选择相同的策略，那么每个参与者得到的收益为 0。参与者 1 知道哪种策略最优（也就是说，他知道"自然"的选择）。三个参与者同时选择策略，求解一个贝叶斯纳什均衡。

附录 1

贝叶斯纳什均衡的正式定义

假设一个博弈涉及参与者 i 同时从行动的可行集合 A_i 中选择行动 a_i。"自然"先行决定每个参与者的类型，这里参与者 i 的类型用 t_i 表示，所有的参与者类型集合用 T_i 表示。每个参与者了解自己的类型（且只了解自己的类型），然后参与者同时选择行动。策略赋予每个参与者每种类型一种选择行动方案。

在选择行动时，一个参与者需要有关于其他参与者类型的信念，定义：

$$t_{-i}=(t_1, t_2, \cdots, t_{i-1}, t_{i+1}, \cdots, t_n)$$

为所有参与者（除了参与者 i）类型的数组，令 T_{-i} 表示 $n-1$ 个参与者的类型空间。$P_i(t_{-i} \mid t_i)$ 表示在知道自己类型的

条件下，参与者 i 的类型分布在 t_{-i} 数组中的概率。如果所有参与者的类型为独立随机变量，那么 $P_i(t_{-i}\mid t_i)$ 的值就不依赖 t_i 的值。但所有参与者的类型可能有关联。这就是 10.3 小节中提到的共同价值和"赢者的诅咒"例子，在这两个例子中，两个投标者的估价都取决于物品的真值，所以她们是正相关的关系。这样的话，

$$\rho_1(s_2\mid s_1)=\begin{cases}\dfrac{1}{2}, & s_2=s_1+4\\ \dfrac{1}{2}, & s_2=s_1-4\end{cases}$$

拥有关于其他参与者类型和策略的信念，一个参与者可以推导其他参与者将会采取什么行动的信念。令 $s_j(\cdot)$ 表示参与者 j 的策略。那么当 $t_j=t'$ 时，$s_j(t')$ 就是参与者 j 的行动。参与者 i 认为参与者 j 选择行动 a' 的概率等于参与者 j 选择 a' 行动类型的概率 —— 也就是说，t_j 的概率取值使得 $s_j(t_j)=a'$。给定关于其他参与者类型和所采取策略的信念，一个参与者可以推导出那个参与者采取什么行动的信念。

给定参与者的类型和选择的行动，令 $V_i(a_1,\cdots,a_n;t_i,t_{-i})$ 表示参与者 i 得到的收益。在许多情况下，一个参与者得到的收益取决于参与者选择的行动和他自己的类型，但这里我们使概率也直接取决于其他参与者的类型。当然，其他参与者的类型间接影响概率，因为另一个参与者的类型决定了选择的行动。

当且仅当，

$$\sum_{t_{-i}\in T_{-i}}V_i(s_1^*(t_1),\cdots,s_n^*(t_n);t_i,t_{-i})\rho(t_{-i}\mid t_i)$$
$$\geqslant\sum_{t_{-i}\in T_{-i}}V_i(s_1^*(t_1),\cdots,s_{i-1}^*(t_{i-1}),a_i,s_{i+1}^*(t_{i+1}),\cdots,s_n^*(t_n);t_i,t_{-i})\rho(t_{-i}\mid t_i),\qquad a_i\in A_i,t_i\in T_i$$

时 $(S_1^*(\cdot),\ldots,S_n^*(\cdot))$ 是一个贝叶斯纳什均衡。

$\sum t_{-i}\in T_{-i}$ 表示 T_{-i} 集合中 t_{-i} 取值得到收益总和 V。对参与者 i 的每个类型来说他的策略 S_i^* 赋予他所选择的行动需要满足条件：使他的期望收益最高，这里的期望收益是相对其他参与者的类型来说的（即他们的行动）。对于所有的参与者，这个条件式都必须成立。

附录 2

类型为连续体时的一级密封价格拍卖博弈

我们已经思考了当投标者的估价都是独立的时的一级密封价格拍卖博弈的简单模型。在附录 2 里，我们假设一个投标者的估价具有无限个数值来丰富这个模型。因为有连续个投标者类型，所以我们需要用到微积分来推导均衡出价策略。

独立私人价值

假设有 $n\geqslant 2$ 个投标者，令 v_i 表示投标者 i 对物品的估价。如果投标者 i 赢得拍卖品，付出的价格为 p，那么他得到的收益为 v_i-p。依据累计均匀分布函数，每个投标者的估价是 [0,1] 之间的任意值：

$$F(v)=\begin{cases}0, & v<0\\ v, & 0\leqslant v\leqslant 1\\ 1, & 1<v\end{cases}\tag{10.1}$$

所以，一个投标者的估价小于或等于 v' 的概率是 $F(v')$，或者是 v'。每个投标者的估价是私人信息，但根据均匀分布，每个投标者的估价在 [0,1] 之间是共同知识。

贝叶斯博弈的过程如下："自然"同时为 n 个投标者选择 n 个估价，且只透露 v_i 的值给投标者 i。然后投标者同时提交出价。赢标者是提交出价最高的投标者，她付给拍卖商的价钱等于她的出价。

思考下面的对称策略组合。其中，估价为 v 的投标者提交出价：

$$b^*(v)=(\frac{n-1}{n})v$$

让我们来证明这是一个对称贝叶斯纳什均衡。换句话说，如果所有的投标者的估价都是均匀分布在 [0,1] 之间的取值，且投标者 j 出价（$\frac{n-1}{n}$）v_j，那么对于投标者 i，出价（$\frac{n-1}{n}$）v_i 为最优策略。对于所有的 v_i 值（也就是投标者 i 的所有类型），这个条件式都必须成立。因为博弈具有对称性，如果对于一个投标者条件式成立，那么对于所有的投标者，条件式都是成立的。

为了验证这个结论，思考投标者 i 的期望收益，因为它的值和投标者的估价和出价相关，我们用 $\pi_j(b_i, v_i)$ 来表示他得到的期望收益。他得到的收益等于他赢得拍卖品的概率和通过赢标获得的收益的乘积，即：

$$\pi_j(b_i, v_i)=\text{Prob}(b_i > b_i, j\neq i)\times(v_i - b_i) \tag{10.2}$$

这里，Prob ($b_i > v_i, j\neq i$) 表示投标者 i 出价高于其他所有投标者出价的概率。（我们不需要考虑他们出价相同的情况，因为不会出现出价相同的情况。）给定投标者 j 的出价策略为（$\frac{n-1}{n}$）v_j，在公式（10.2）的概率表达式中，我们可以把 b_j 替换为（$\frac{n-1}{n}$）v_j：

$$\text{Prob}(b_i > b_j, j\neq i)=\text{Prob}(b_i > (\frac{n-1}{n})v_j, j\neq i)$$
$$=\text{Prob}((\frac{n-1}{n})b_i > v_j, j\neq i)$$

因为所有投标者的出价都是不同的取值，所以我们得到：

$$\begin{aligned}&\text{Prob}((\frac{n-1}{n})b_i > v_j)\\&=\text{Prob}((\frac{n}{n-1})b_i > v_1)\times\cdots\times\text{Prob}((\frac{n}{n-1})b_i > v_{i-1})\times\\&\text{Prob}((\frac{n}{n-1})b_i > v_{i+1})\times\cdots\times\text{Prob}((\frac{n}{n-1})b_i > v_n)\end{aligned} \tag{10.3}$$

使用公式（10.1）的函数表达式，我们得到：

$$\text{Prob}((\frac{n}{n-1})b_i > v_j)=(\frac{n}{n-1})b_i$$

把这个结果代入公式（10.3）中，得到

$$\text{Prob}((\frac{n}{n-1})b_i > v_j, j\neq i)=\left[(\frac{n}{n-1})bi\right]^{n-1}$$

最后，把这个表达式代入公式（10.2），给定其他投标者采取的策略和他们的估价概率分布，我们推导出投标者 i 得到的期望收益为：

$$\pi_i(b_i, v_i)=\left[(\frac{n}{n-1})bi\right]^{n-1}(v_i - b_i)$$

为了求解投标者 i 的最优出价策略，我们只需要条件式的第一项满足：

$$\frac{\partial\pi_i(b_i,v_i)}{\partial b_i}=-\left[\left(\frac{n}{n-1}\right)b_i\right]^{n-1}+(n-1)\left(\frac{n}{n-1}\right)\left[\left(\frac{n}{n-1}\right)b_i\right]^{n-2}(v_i-b_i)=0 \tag{10.4}$$

对投标者 i 来说，最优出价策略就是满足公式（10.4）条件式 b_i 的值。简化重新整理公式（10.4），我们可以求出 b_i：

$$b_i=(\frac{n-1}{n})v_i$$

这正是我们想要证明的。如果所有的投标者使用相同的出价策略，那么对于投标者 i 而言，使用这个策略也是最优策略。

然后每个投标者成比例地隐藏自己的出价的 $\frac{1}{n}$。很明显，投标者不愿意给拍卖商的出价等于她的估价，因为这样即使她赢得拍卖品，她付给拍卖商的钱等于她对物品的估价，那么她赢没赢所得的收益都为 0。而出价低于她的估价时，或者她输了，得到的收益为 0；或者她赢了，得到的收益为正值。于是，如果她的出价低于她的估价，

那么她的期望收益为正值，高于出价等于估价的情况。问题是投标者如何隐藏自己的出价。出价越低，得到的收益越高，但获胜的概率越小。每个投标者的均衡策略是：提交的出价等于估价的 $\frac{n-1}{n}$。因为 $\frac{n-1}{n}$ 的值随着 n 的值增大而增大，所以拍卖中投标者越多，投标者的出价就越高。这是有道理的，因为更多的投标者意味着更激烈的竞争，这样每个投标者隐藏的出价就不能太高，以保证在合理的范围赢得拍卖品。

共同价值

假设有 $n \geqslant 2$ 个投标者。对于所有的投标者，正在拍卖的物品的真正价值为 v。每个投标者得到一个不准确的观测信号 v，v 是根据均匀分布由“自然”选择的处在 [0,1] 之间的一个任意值。投标者 i 观测信号 s_i 的累计分布函数为：

$$F(S_i)=\begin{cases}0, & s_i<0\\ s_i, & 0\leqslant s_i\leqslant 1\\ 1, & 1<s_i\end{cases}$$

只有投标者 i 知道自己的观测信号；所以，一个投标者的观测信号就是他的类型，且类型空间为 [0,1]。每个投标者的观测信号都是 [0,1] 之间的任意一个值是共同知识。最后，假设“自然”随机决定物品的真正价值，因此假设它等于所有投标者观测信号的平均值：

$$v=\left(\frac{1}{n}\right)\sum_{j=1}^{n}s_j \tag{10.5}$$

投标者参加的是一级密封价格拍卖。这就是说，如果投标者 i 赢标，那么他得到的收益为 $v-b_i$。这里 b_i 是他的出价，尽管在他赢得拍卖品之后才能知道 v 的值。

在推导一个贝叶斯纳什均衡时，让我们推测它和投标者的观测信号呈线性相关。换句话说，存在一个 $\alpha>0$ 的值使得：

$$b_j=\alpha s_j \tag{10.6}$$

在他提交的出价为最高的条件下，投标者 i 的期望收益等于他赢标的概率（也就是说，他的出价比其他投标者都高）乘以他的期望收益：

$$\text{Prob}\,(b_i>b_j,\ j\neq i)\times\{E[v\mid s_i, b_i>b_j,\ j\neq i]-b_i\} \tag{10.7}$$

$E[v\mid s_i, b_i>b_j,\ j\neq i]$ 是投标者 i 的期望价值，这不仅和他的观测信号有关，还和他是否提交的是最高价有关。后者说明了其他投标者的观测信号情况，所以也说明了物品的真值。

现在我们假设其他的投标者都使用公式（10.6）中的出价策略，把公式（10.7）中的 b_j 替换为 αs_j：

$$\begin{aligned}&\text{prob}\,(b_i>\alpha s_j,\ j\neq i)\times\{E[v\mid s_i,b_i>\alpha s_j,\ j\neq i]-b_i\}\\&=\text{prob}\left(\frac{b_i}{\alpha}>s_j,\ j\neq i\right)\times\left\{E\left[v\mid s_i,\frac{b_i}{\alpha}>s_j,\ j\neq i\right]-b_i\right\}.\end{aligned}$$

接下来，替换公式（10.5）中 v 的表达式：

$$\begin{aligned}&\text{prob}\left(\frac{b_i}{\alpha}>s_j,\ j\neq i\right)\times\left\{E\left[\left(\frac{1}{n}\right)\left(s_i+\sum_{j\neq i}s_j\right)s_i,\frac{b_i}{\alpha}>s_j,\ j\neq i\right]-b_i\right\}\\&=\text{prob}\left(\frac{b_i}{\alpha}>s_j,\ j\neq i\right)\times\left\{\left(\frac{s_i}{n}\right)+\left(\frac{1}{n}\right)E\left[\sum_{j\neq i}s_j\middle|\frac{b_i}{\alpha}>s_j,\ j\neq i\right]-b_i\right\}\\&=\text{prob}\left(\frac{b_i}{\alpha}>s_j,\ j\neq i\right)\times\left\{\left(\frac{s_i}{n}\right)+\left(\frac{1}{n}\right)\sum_{j\neq i}E\left[s_j\middle|\frac{b_i}{\alpha}>s_j\right]-b_i\right\}.\end{aligned}$$

第二行是由于投标者 i 知道 s_i 的值但不知道 s_j 的值得来，所以 $E[s_i]=s_i$。第三行是因为所有的观测信号都是互不相关的随机变量。利用 s_j 的均匀分布，我们得出投标者 i 的期望收益为：

$$\left(\frac{b_i}{\alpha}\right)^{n-1}\left[\left(\frac{s_i}{n}\right)+\left(\frac{n-1}{n}\right)\left(\frac{b_i}{2\alpha}\right)-b_i\right] \tag{10.8}$$

其中：

$$E\left[s_j \middle| \frac{b_i}{\alpha}>s_j\right]=\frac{b_i}{2\alpha}$$

投标者 i 选择策略 b_i 使得公式（10.8）的值最大。条件式的第一项是：

$$\frac{\partial\bullet}{\partial b_i}=(n-1)\left(\frac{1}{\alpha}\right)\left(\frac{b_i}{\alpha}\right)^{n-2}\left[\left(\frac{s_i}{n}\right)+\left(\frac{n-1}{n}\right)\left(\frac{b_i}{2\alpha}\right)-b_i\right]+\left(\frac{b_i}{\alpha}\right)^{n-1}\left(\frac{n-1-2\alpha n}{2\alpha n}\right)=0$$

解关于 b_i 的等式，我们得到：

$$b_i=\left(\frac{2\alpha}{2\alpha n-(n-1)}\right)\left(\frac{n-1}{n}\right)s_i \tag{10.9}$$

记得我们推测过对称均衡出价策略是 $b_i=\alpha s_i$。方程公式（10.9）确实是线性方程，并且通过将 s_i 替换为协同因素 α 的表达式，我们可以解出：

$$\alpha=\left(\frac{2\alpha}{2\alpha n-(n-1)}\right)\left(\frac{n-1}{n}\right)$$

求解关于 α 的方程式，我们得到：

$$\alpha=\frac{(n+2)(n-1)}{2n^2}$$

总之，一个对称贝叶斯纳什均衡策略使得一个投标者使用的出价策略为：

$$b_i=\left(\frac{n+2}{2n}\right)\left(\frac{n-1}{n}\right)s_i$$



第 11 章

你的行为暴露了你的身份：信号传递博弈

学习目标

11.1 精炼贝叶斯纳什均衡

11.2 求解信号传递博弈

本章导读

掩盖的事，没有不露出来的；隐藏的事，没有不被人知道的。

—《英皇钦定版圣经：路加福音》，第 12 章

在许多的策略式博弈中，一个参与者知道另一方想了解的信息。内维尔·张伯伦在和阿道夫·希特勒谈判时，就想了解希特勒的真实意图。怀亚特·厄普在决定是否拿枪时，也想弄清他所面对的陌生人有什么样的本事。当进行拍卖投标时，一个投标人可以通过掌握其他竞标者对拍品的估价，从而知晓他们的竞拍价格。一个参与者可以就她所知的信息做出最佳决策，但同时她也想了解有关博弈另一方的信息。

在一些情形中，一个参与者有机会获取另一参与者知道的信息。例如，假定希特勒不确定张伯伦愿意参战的收益情况。虽然希特勒无法了解张伯伦的心思，但是他可以通过观察张伯伦是否妥协来判断张伯伦的收益。或者分析另一个汽车经销商与顾客之间讨价还价的情形。尽管卖方不知道买方想要购车的急切程度，可是他可以从买方提供的购买价格中做出推测。为什么汽车经销商会经常向你询问你愿意购买的价格？因为他们想先让你透露出一些有关你购车意愿的信息，然后根据你提出的购买价格来表明他们愿意卖出的价格。

在本章以及下一章当中，我们会模拟和分析这些情形。这些情形中出现的基本模型叫做**信号传递博弈**（signaling game），该博弈有两个参与者：发送者和接收者。发送者的类型是属于她的私人信息，因而接收者并不知晓。发送者选择某一行动作为向接收者发出的信息。接收者观察发送者的行动之后，做出自己的行动选择。然而，接收者采取的行动是否最优取决于发送者的类型（也取决于发送者的行动）。

为了解信号传递博弈中会出现怎样的结果，不妨先假设一个你在毕业时肯定会碰到的情形。某人接受了一个管理培训生的职位。经理将根据这名培训生在观察期的表现来决定是否长期雇用他。在培训期间，考察的内容之一就是求职者愿意努力工作的程度。为了便于分析，假设有两种类型的员工：懒散型和勤勉型。懒散的员工倾向于每周工作 40 个小时（即标准

PDG

工作日），而勤勉的员工愿意将工作时间延长至 60 个小时。

假定经理只想雇用一名勤勉的员工。当然，无论是懒散的员工还是勤勉的员工都想被长期雇用。所以如果一个事实上是懒散的人该怎么做才能被经理留下呢？虽然她不愿意每周工作 60 个小时，但如果这么做会让经理觉得她是一个勤勉的员工进而被长期聘用的话，那么她也许会在培训期间增加自己的工作时间。现在假设这名懒散的员工确实延长了工作时间。此时，经理（如果智商不低于下属的话）应该可以看出一个工作 60 个小时的管理培训生不一定是勤勉的，因为他可能是一个伪装成勤勉员工的懒散的人。这样的话，这名经理就不能从培训期间员工的努力程度来判断员工所属的类型。

思考一下勤勉员工的情况。如果她够聪明的话，就会意识到即使是懒散的员工也会为了掩盖自己的真实类型而勤勉工作。那么，勤勉员工该怎么做呢？将每周工作时间增加至 80 个小时！勤勉员工可能会以进一步增加工作时间的极端方式将自己和懒散员工区别开来，从而使经理弄清究竟谁应该留下。

对经理来说，从每周 80 小时的工作时间推断出该员工确实是勤勉的关键在于，懒散员工是不愿意工作 80 小时的，即使只有那样才有被雇用的机会。

信号传递博弈往往涵盖许多微妙的战略。本例中，一个懒散的实习生想尽办法千方百计地愚弄经理。勤奋的实习生则想要和那些懒散的人划清界限，而经理要做的是把这两类人区别开来，以便决定是否继续留用。

本章的主要目标是学习如何求出信号传递博弈的解，进而对行为作出准确预测。在下一节当中，我们先回顾适用于信号传递博弈的一个求解方法，然后将这一方法应用于管理培训生博弈的情形。在第 3 节中，我们会分析涉及信号传递博弈的各种情形，包括二手车市场，一触即发的核战争和恋爱。

11.1 精炼贝叶斯纳什均衡

一个信号传递博弈中存在三个阶段：

阶段 1：“自然”决定发送者的类型。
阶段 2：发送者根据自己的类型选择某一行动。
阶段 3：接收者观察发送者的行动，依据这一新信息更新自己对发送者类型的信念后选择一个行动。

发送者的策略是根据每个可能的类型选择一个行动，接收者的策略是根据发送者每个可能的行动选择一个行动。求解这一博弈的方法被称为**精炼贝叶斯纳什均衡**。之所以如此定义是因为精炼贝叶斯纳什均衡是对贝叶斯纳什均衡的精炼，就如同子博弈精炼纳什均衡是对纳什均衡的精炼一样（这颇有点学术能力评估考试中一个问题答案的味道）。

精炼贝叶斯纳什均衡中有两个关键的概念：序贯理性和一致信念。**序贯理性**（sequential rationality）是指在博弈的每一阶段，参与者的策略会基于对其他参与者行动的信念指定一个最佳行动。正如同子博弈精炼纳什均衡中所表述的那样，无论参与者处于博弈中的哪一阶段，都会根据她的信念采取最大化自己期待收益的行动。在信号传递博弈中，序贯理性要求发送者所选的策略是她可能类型中的最佳策略（这一点与贝叶斯纳什均衡类似），接收者所选的策略是应对发送者每个可能行动的最佳策略。

需要提醒的是序贯理性要求的是根据信念采取最佳行动。正如你所想的，信念不可能是已有的现成的东西，而是基于对其他参与者的策略行为巧妙推断而得出的。在信号传递博弈中，接收者起初就有关于发送者类型的信念集合，该信念被称为**先验信念**（prior belief）并且是“自然”赋予的可行类型；之后他通过观察发送者的行动确定自己的行动。因为发送者的行动可能包含发送者所属类型的信息，所以接收者会更新自己原有的信念以推导出一套**后验信念**（posterior belief）或者是基于发送者行动的信念。如果接收者的后验信念与发送者最大化自身利益的行动相一致，那么他就有**一致信念**(consistent beliefs)。换句话说，接收者会自问：“既然已经观察了发送者的行动，那么什么类型的发送者会采取这样的行动呢？”

为了使一致信念的概念更具直观性，图 11—1 列举

了一个发送者三种不同类型的策略。在这个例子当中，有四种类型的发送者，分别是贪婪型（greedy）、节俭型（frugal）、吝啬型（miserly）和慷慨型（generous）。接收者的先验信念认为发送者是贪婪型的概率为0.2，节俭型的概率为0.4，吝啬型的概率为0.1，以及慷慨型的概率为0.3。假设发送者的可行行动是 A, B, C, D, E 和 F。

从图11—1（a）所表示的策略可以看出，贪婪型选择行动 B，节俭型选择行动 D，吝啬型选择行动 A，慷慨型选择行动 E。在已经知道（或猜想）发送者的策略，以及观察到发送者的行动的条件下，接收者会做出什么推断呢？倘若接收者看到的是行动 A，那么他知道发送者一定属于吝啬型，因为根据发送者所选的策略，只有吝啬型会选择行动 A。类似的，如果接收者见到的是行动 B，那么发送者一定属于贪婪型，以此类推。因为每一类型选择的行动不同，所以从观察到的行动就可以判断出发送者的类型。这种情况就属于**分离策略**(separating strategy)，即每一类型的参与者具有不同行动的策略。因而，接收者可以从她所看到的活动中区分出每个参与者的类型。

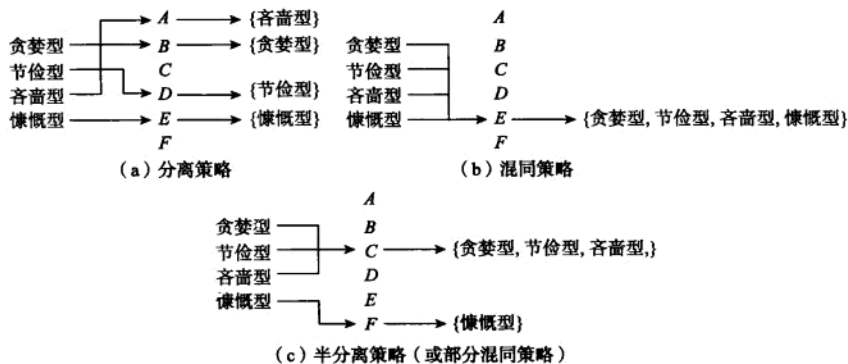


图11—1　发送者三种不同类型的策略

然而，分析图11—2b中的**混同策略**(pooling strategy)之后就会发现，行动并不总是透露实情的。在这种情形下，不论发送者的类型是什么，她所选择的策略指定采取相同的行动，即不管她属于贪婪型、节俭型、吝啬型还是慷慨型，她都会选择行动 E。接收者就无法从发送者的行动判定她的类型。借助混同策略，无论发送者的真实类型是什么，所有的发送者类型都选择相同的行动。

图11—1c表述的是介于分离策略和混同策略之间的情形。无论是贪婪型、节俭型还是吝啬型的发送者都会选择行动 C，只有当她属于慷慨型时才会选择行动 F。若接收者看到的是行动 F，则他推断发送者是慷慨型。倘若接收者看到的是行动 C，那么除了可以排除发送者是慷慨型之外，无法做出其他的推测，这样接收者对发送者的判断就可能是贪婪型、节俭型或是吝啬型。发送者的行动对她所属类型只透露出部分而不是全部的信息。这种策略就属于**半分离策略**（semiseparating strategy）或**半混同策略**(semipooling strategy)，该策略既不同于每个类型选择不同行动的策略也不同于所有类型选择相同行动的策略。

一致信念至少要求参与者的后验信念赋予采取不同行动类型的概率为零。可问题是，所剩类型采取相同行动的概率又是多少。在分离策略中，分离策略基本上不涉及概率分配问题，如图11—1a所示，因为所有与所观察行动不一致的类型被剔除之后就剩下了一种类型。在这种情形下，接收者认为这一类型的概率是1。例如，接收者观察到的行动是 B 的话，那么接收者会认为发送者是贪婪型的概率为1。

如图11—1b所描述的那样，发送者使用混同策略时也可以直接推出一致信念。因为发送者的行动没有提供任何信息（所有类型都选择相同行动），所以接收者的先验信念和后验信念是一致的，即“自然”赋予各种类型发送者的概率。

半分离策略是这些策略中最不容易计算出概率的策略，如图11—1c所示。如果接收者看到发送者选择行动 C，那么他推断发送者是慷慨型的概率为零。但是他如何推

算贪婪型、吝啬型和勤俭型的发送者会选择行动 C 的概率呢？在回答这个问题之前，不妨先引入由 18 世纪神学家托马斯·贝叶斯提出的贝叶斯法则。

假设接收者起初判断发送者是可行类型 $\{1, 2, \cdots, n\}$ 中的任意一个，并且起先认为发送者是类型 i 的概率是 p_i。接收者之后得知发送者的类型是集合 $\{1, 2, \cdots, n\}$ 的子集合 T 中的一个类型。在上述的例子中，子集合 T 就是｛贪婪型，节俭型，吝啬型｝，因为接收者认识到发送者不是慷慨型。对于每个不属于 T 的类型，参与者的后验信念认为他们的概率为零。对于每个属于 T 的类型，贝叶斯法则得出后验概率等于先验概率 p_i 除以子集合 T 中所有类型概率的总和。在本例中，概率总和是 0.7 (= 0.2+0.4+0.1)，所以发送者是贪婪型的后验概率是 $\frac{2}{7}$，节俭型的后验概率是 $\frac{4}{7}$，吝啬型的后验概率是 $\frac{1}{7}$。可以看出贝叶斯法则使得相对概率保持不变。比如，按照先验信念，发送者是节俭型的概率是贪婪型概率的两倍（0.4 对 0.2）。得知发送者不是慷慨型后，接收者赋予贪婪型和节俭型的概率分别是 $\frac{2}{7}$ 和 $\frac{4}{7}$，所以发送者是节俭型的概率仍是贪婪型的两倍。这一结论是有道理的，因为接收者了解的所有信息只是发送者不是慷慨型，他对发送者是贪婪型、节俭型或吝啬型的相对概率一无所知。

因为序贯理性规定接收者所选的策略是应对发送者采取任何行动的最佳策略，所以有必要明确接收者对发送者每个行动的后验信念，从而使不确定的情形通过分析图 11—1a 中的策略而变得清晰起来。如果接收者观察到发送者的行动是 F，那么他的一致信念是什么？问题是根据接收者对发送者所选策略的信念，没有一个类型的发送者会选择行动 F。接收者无法给所有发送者类型赋予零概率，因为发送者必须是其中的某一个类型。在这种状况下，一致性要求对信念没有任何约束力。我们对这种“意外情况”描述的信念是任意的，因为在这一问题上的我们没有什么可以遵循的准则。

总的来说，信号传递博弈中是否出现精炼贝叶斯纳什均衡取决于发送者的策略，接收者的策略和满足以下条件的接收者的信念集合（根据对发送者行动的观察，推断出发送者类型的信念）：

- 对于每个类型的发送者，他的策略要求针对接收者做出的反应，采取最大化自身期待收益的行动。
- 对于发送者的每个行动，接收者的策略要求接收者根据对发送者类型的信念采取最大化自己期待收益的行动。
- 接收者对发送者类型的信念（根据对发送者行动的观察）与发送者的信念和贝叶斯法则相一致。这一条件只适用于根据发送者的策略，接收者对某一类型采取的行动。

案例研究

管理培训生

下面我们就将这一求解概念应用于之前的管理培训生博弈中。开始之前，需要明确相关收益和信念。假定一个实习生被雇用得到的收益是 130，而不被聘用的收益是 70。就努力程度而言，每个实习生有三种选择：40、60 和 80 个小时工作时间，而个人的损失则由她的类型决定，如表 11—1 所示。

表 11—1　不同努力程度的相关损失情况

类型	40个小时	60个小时	80个小时
懒散型	50	75	120
勤勉型	30	50	80

实习生的收益等于被雇用（或不被雇用）的价值减去个人努力的损失。例如，如果懒散型每周工作 60 个小时并且得到这一职位的话，那么她的收益是 55 (=130 − 75)，而她没有得到职位的收益是 −5 (=70 − 75)。经理雇用一个勤勉员工的收益是 100，雇用懒散员工的收益是 25，选择不雇用的收益是 60。这意味着他宁愿空着职位也不愿雇用懒散员工。

在观察实习生的努力程度之前，先假定经理认为实习生是懒散型的概率为 0.75，是勤勉型的概率为 0.25。这些信念意味着经理只有在获取更多有关实习生类型的信息之后才会选择雇用她，那么雇用她的期望收益是 $0.75 \times 25 + 0.25 \times 100 = 43.75$，而这一收益小于使职位空缺得到的收益 60。最后，注意观察实习生对每个可行类型（懒散型或勤勉型）采取不同行动（工作 40, 60 或 80 个小时）的策略，以及经理对实习生的每个可行行动（工作 40, 60 或 80 个小时）采取不同行动（雇用或不雇用）的策略。这个贝叶斯博弈的情形如图 11—2 所示。

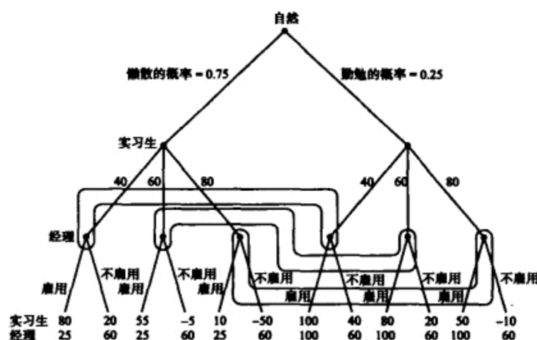


图 11—2 管理实习生博弈

要学会如何应用精炼贝叶斯纳什均衡求出博弈的解，不妨先分析下面的策略和信念集合：

- 实习生的策略：
  若选择懒散，则工作 40 个小时。
  若选择勤勉，则工作 80 个小时。
- 经理的策略：
  若实习生工作 40 或 60 个小时，则不雇用她。
  若实习生工作 80 个小时，则雇用她。
- 经理的信念：
  若实习生工作 40 个小时，则认为她懒散的概率为 1。
  若实习生工作 60 个小时，则认为她懒散的概率是 0.6，勤勉的概率是 0.4。
  若实习生工作 80 个小时，则认为她勤勉的概率为 1。

要判断上述情形是不是精炼贝叶斯纳什均衡，不妨先从实习生的策略着手。给定经理的策略，实习生采取各种行动得到的收益如表 11—2 所示。比如，实习生属于懒散型，并工作 60 个小时的话，那么她的收益是—5，因为这种情况下她的个人损失是 75，而且不被雇用（按照经理的策略推断）。假定实习生属于懒散型，那么她工作 40 个小时是最佳策略，因为这样做得到的收益 20 大于工作 60 个小时的收益 –5 和工作 80 个小时的收益 10。可是，如果实习生属于勤勉型，那么她就应该工作 80 个小时。这么做她得到的收益是 50（因为她被雇用），但如果工作时间越少得到的收益就更少，因为她不被雇用。因而，给定经理的策略，实习生对两种类型选择的策略都是最佳策略。

表 11—2 实习生的收益（根据经理的策略）

类型	40个小时	60个小时	80个小时
懒散型	20	–5	10
勤勉型	40	20	50

在判断经理的最优策略之前，我们先证明一下他的信念是一致的。实习生工作 40 个小时的话，经理的信念认为她懒散的概率是 1，这和实习生的策略相一致，因为只有懒散的实习生工作 40 个小时。如果他观察实习生工作 80 个小时，那么他认为她勤勉的概率是 1，因为根据实习生的策略，只有勤勉型实习生工作 80 个小时。最后，倘若观察到实习生工作 60 个小时，则信念不定，因为这一行动和实习生的策略不一致。我们任意假定经理认为实习生是懒散型的概率为 60%。

接下来就是用这些一致信念去推断经理的策略是否是序贯理性的。思考实习生采取的每个可行行动。实习生工作 40 个小时的话（假定经理肯定认为她是懒散的），那么经理雇用她的收益是 25，不雇用她的收益是 60，所以按照自己策略的指定，经理会选择不雇用她。相反，如果实习生工作 60 个小时，那么根据经理的后验信念，他雇用实习生的期望收益是 $0.6 \times 25+0.4 \times 100=55$，而不雇用她的期望收益是 60。因而，在实习生工作 40 或 60 个小时的条件下，经理选择不雇用她是最佳策略。最后，实习生工作 80 个小时的话，那么经理认为她是勤勉的，在这种情形下选择雇用她（得到 100 的收益）的策略要优于选择不雇用她（得到 60 的收益）的策略。这样，经理对实习生每个可行行动的策略都是最佳策略。由此得出，这一情形就是精炼贝叶斯纳什均衡。

11.1 理解测验

再分析之前的策略组合，但假定当经理观察到实习生工作60个小时，他给她是懒散型赋予的概率是p。（在之前的分析中，我们设定p=0.6）。求出p的值，从而使这一策略组合和信念集合成为一个精炼贝叶斯纳什均衡。

注意在这个例子中讨论的精炼贝叶斯纳什均衡是一个分离均衡，因为可以通过发送者的行动推断出她的类型。工作 80 个小时表明实习生是勤勉的，而工作 40 个

小时则意味着她是懒散的。另一个分离策略是勤勉型工作 60 个小时，而懒散型工作 40 个小时，这一情形将在理解测验 11.2 中得到验证。然而，这一情形却不是一个精炼贝叶斯纳什均衡，因为懒散型会选择工作 60 个小时，伪装成一个勤勉员工从而达到被雇用的目的。这样的话，为了有效表明自己的类型，勤勉的实习生就需要工作 80 个小时。


点拨

为了有效表明自己的类型，一个优秀类型的发送者（如勤勉的）可能需要改变自己的行动，从而和模仿自己的不太优秀类型的发送者区别开来。

图 11—2 所描述的是无论实习生的类型是什么，她的工作时间都相同的情形是一个混同均衡。分析下面的策略组合和信念集合：

■ 实习生的策略：不管懒散还是勤勉，工作 40 个小时。

■ 经理的策略：不雇用她（无论她多么努力工作）。

■ 经理的信念：给懒散的实习生赋予 0.75 的概率（不论她多么努力工作）。

既然经理在任何情况下都不会雇用自己，那么实习生选择工作 40 个小时的策略是最佳策略。注意经理的后验信念和先验信念是一致的。当他观察到实习生工作 40 个小时，这些信念与实习生的策略一致，因为两个类型的实习生都只会工作 40 个小时。由于实习生的行动提供不了任何信息，信念保持不变。当实习生工作 60 或 80 个小时，信念就不确定，我们任意假定和先验信念相同。最后一个问题是经理策略的序贯理性。根据他的信念，当且仅当：

$$60 \geqslant 0.75 \times 25 + 0.25 \times 100 = 43.75$$

时，他选择不雇用实习生的策略是最优的。不等式的左边是不雇用她的收益，右边是雇用她的期待收益。因而，两个类型的实习生都选择最低工作时间的策略也是一个均衡。

11.2 理解测验

思考一个策略组合，其中勤勉型工作60个小时，懒散型工作40个小时。如果实习生工作至少60个小时，经理就雇用她，而如果她工作40个小时就不雇用她。若实习生工作40个小时，则经理认为她是懒散型的概率为1，而她工作60个或80个小时的话，经理认为她是懒散型的概率为零。试证明这一情形不是一个精炼贝叶斯纳什均衡。

11.2 求解信号传递博弈

在本章的后续部分中，我们将对一些难度递增的例子予以评析。次品博弈研究的是二手车市场，其中发送者（车主）有三种可行的类型（对应于真实的车况）。恋爱博弈中的发送者和接收者都可以有私人信息。最后，边缘政策博弈分析的是发送者使用混合策略的情况。通过对这些案例的分析，我们不仅可以学会如何求出信号传递博弈的解，还能得到一些启示，即在没有信息的情况下如何做出富有策略性的行动。

○案例研究

次品与二手车市场博弈

在任何一个经济体中，私人信息存在的一个典型情境就是某一产品（如汽车或房子）的卖方和潜在买方之间的博弈情形。虽然买方能够从外观判断这一产品的质量，但是这种判断方式常常检查不出只有卖方才知道的产品问题。所以对于一个打算买车的人来说，一个很大的顾虑是卖方之所以选择卖掉这辆车是因为她知道这车是个次品。如果的确如此，那么买方就可能买到了一辆劣质车，而之所以出现这一状况不是由于偶然，而是卖方精心设局的结果。但关键问题是卖方和买方之间信息的不对称是否会使得双方共赢的交易无法产生。下面我们就运用一些博弈理论来深入研究二手车市场的运行机制。

假设待售汽车的质量有三种可行类型，分别是高品质，中等品质和低品质。卖方知道汽车的真实质量状况，

① 记住买方是无法观察出汽车质量的。之后我们分析某一年代的某一款车型，买方不知道的信息是这一汽车是否存在如变速器或刹车等的质量问题。——作者注

而买方不知道，这样卖方的类型就是汽车质量的类型。[①]买方开始认为这车是高品质的概率为 20%，中等品质的概率为 50%，低品质的概率是 30%。根据这些概率判断，“自然”选择卖方的类型；然后卖方决定是否出售这辆车，如果出售，定价是多少（可以是任何一个正整数）。如果出售这辆车的话，那么买方看到价格之后就会决定是否购买。为了简单起见，通常出现的讨价还价方式不予考虑，所以卖方要么选择原价出售要么选择不出售。这样卖方的策略是给汽车质量的每个类型制定不同的价格。买方的策略是选择接受或不接受每个可行的价格（当然在汽车要出售的前提下）。

如果卖掉汽车，卖方的收益是获得的汽车价款。如果汽车没有被卖掉，卖方的收益是他赋予汽车的价值，如表 11—3 所示。倘若买方决定购车，那么他的收益是汽车对他产生的价值减去他支付的汽车价格。汽车对买方的价值取决于汽车真正的质量状况，而这一点只能在购买之后驾车的过程中才能被发现。因而，买方在购车时并不知道汽车对他产生的价值。如果他不买车的话，那么收益为零。

表 11—3　二手车市场

质量	概率	卖方的估值	买方的估值
高品质	0.20	20 000	24 000
中等品质	0.50	15 000	18 000
低品质	0.30	10 000	12 000

表 11—3 显示出具备完全信息时，总是存在出售汽车的概率，因为买方对汽车的估值高于卖方的估值。汽车为高品质时，买方的估价是 24 000，而卖方的估价只是 20 000；汽车的质量处于中等品质时，买方的估值比卖方的估值高 3 000（=18 000－15 000）；汽车为低品质时，买方的估值比卖方的估值高 2 000（=12 000－10 000）。问题是买方和卖方之间信息的不对称是否会影响交易成交。

由此自然想到的一个问题是高品质的汽车是否销售量更大。例如，分析一个分离策略中，汽车的质量为高品质时，卖方制定的价格 p^h，质量为中等品质时的定价 p^m 和质量为低品质时的定价 p^l，并且 $p^h > p^m > p^l$。根据卖方的这一策略，买方的一致信念会在他看到价格 p^h 时，认为汽车是高品质的概率为 1；当看到价格 p^m 时，认为汽车是中等品质的概率为 1；当看到价格 p^l 时，认为汽车是低品质的概率为 1。若 $p^h \leq 24\,000$（这样推断汽车是高品质的价格就不会高于高品质汽车对买方的价值），则买方愿意以较高的价格购买；若 $p^m \leq 18\,000$，则以中等价格购买；若 $p^l \leq 12\,000$，则以较低的价格购买。比如，当 $p^h = 22\,000$，$p^m = 17\,000$ 和 $p^l = 10\,500$ 时，买方就愿意以相应的价格购买。

现在我们面临一个两难问题：卖方的策略不是最佳策略。卖方以 10 500 的价格（暗示质量较差）出售一辆低品质汽车时得到的收益是 10 500，但假如她把定价改为 22 000，这样就可以误导买方认为这是辆高品质的汽车，从而得到 22 000 的收益。

刚才说明的是高品质汽车的销售量并不会增大。道理很简单：如果高品质汽车的销售量更大的话，那么拥有低品质汽车的卖方就会模仿拥有高品质汽车卖方的定价行为，从而误导买方以更高的价格购买。拥有高品质汽车的卖方可以制定较高的价格，那么低品质汽车的买方同样可以轻松地制定较高的销售价格。我们从中得到的启示是在二手车市场，价格并不是质量可靠的保证。

另外一种可能性是所有的汽车售价相同。为了分析这一情形，不妨先思考一个混同策略，其中无论汽车是低品质、中等品质还是高品质，卖方都制定相同的价格（用 $\bar{p}$ 表示）[①]：

- 卖方的策略：
  不管汽车是低品质、中等品质还是高品质，都以价格 $\bar{p}$ 出售。
- 买方的策略：
  若 $p \leq \bar{p}$，则购买汽车。
  若 $p > \bar{p}$，则不买汽车。
- 买方的信念：
  不论价格是多少，赋予低品质汽车 0.3 的概率，中等品质汽车 0.5 的概率，高品质汽车 0.2 的概率。

目标是求出一个二手车的价格，即 $\bar{p}$ 的值，从而使得之前的情形成为一个均衡。先从对买方信念的分析开始，只有当观察到的价格是 $\bar{p}$ 时，才能应用一致要求，因为根据卖方的策略，这一价格是唯一采纳的价格。而既然不论汽车的质量如何，卖方都会将价格定为 $\bar{p}$，那么买方的

① 显然，如果对于 $\bar{p}$ 中的某一价值而言，这不是一个均衡的话，那么也就不存在混同均衡。——作者注

后验信念就一定会和他的先验信念相一致。因而，这些信念就是一致的。若 $p \neq \overline{p}$，则信念不定，这时我们任意假定和买方的先验信念相同。这样的话，买方就无法从卖方的定价中了解到任何有关汽车质量的信息。

现在将注意力转向买方的策略，我们推出当且仅当：

$$0.2 \times (24\,000 - \overline{p}) + 0.5 \times (18\,000 - \overline{p}) + 0.3(12\,000 - \overline{p}) \geq 0 \quad (11.1)$$

时，以价格 $\overline{p}$ 买车是一个最佳策略。譬如，因为预期的概率是 0.2，所以汽车被认为是高品质的，这样买方对汽车的估计是 24 000，得到的收益是 $24\,000 - \overline{p}$。通过求公式（11.1）中 $\overline{p}$ 的解，可以得出 $\overline{p} \leq 17\,400$。也就是说，对买方而言，汽车的期望价值（即 17 400）必须至少和她支付的价格相同。因而，如果这一情形是一个均衡的话，那么卖方的定价不能超过 17 400，否则买方不会购买。

最后，我们需要推断卖方的策略是否是最佳策略。按照卖方的策略，他得到的收益是 $\overline{p}$。对于每个卖方类型，这一收益必须至少和卖方对汽车的估价，即：

$$\overline{p} \geq 20\,000，\overline{p} \geq 15\,000 \text{ 和 } \overline{p} \geq 10\,000$$

相等才可以。上述不等式的顺序分别对应于汽车是高品质、中等品质和低品质的情形。所以 $\overline{p}$ 必须至少等于 20 000。

可至此产生一个问题：汽车是高品质时，卖方愿意出售的价格必须至少是 20 000，而买方无法确定汽车是低品质、中等品质还是高品质时，愿意购买的价格不能超过 17 400。就算你没有数学博士学位也知道没有一个数是既大于 20 000 又小于 17 400 的。所以得出的结论是不存在一个价格 $\overline{p}$ 可以使所有汽车质量类型以相同价格出售的情形，因而该情形也不是一个精炼贝叶斯纳什均衡。

总之，汽车不能以不同的价格销售（我们的第一个结论），也不能以相同的价格销售（我们的第二个结论）。那么还剩下什么？回想卖方还可以选择不出售汽车。所以分析一下卖方的半分离策略和以下情形是否是精炼贝叶斯纳什均衡。

■ 卖方的策略：

若汽车是低品质或中等品质的，则定价为 $\overline{p}$。

若汽车是高品质的，则不出售。

■ 买方的策略：

若 $p \leq \overline{p}$，则购买汽车。

若 $p > \overline{p}$，则不买汽车。

■ 买方的信念①：

若 $p \leq \overline{p}$，则认为汽车是低品质的概率为 0.375，中等品质的概率为 0.625，高品质的概率为 0。

若 $p > \overline{p}$，则认为汽车是低品质的概率为 1。

这次的分析先从卖方开始。当且仅当 $\overline{p} \geq 15\,000$ 时，有中等品质汽车的卖方才会认为卖掉汽车是最佳策略，因为只有这时他得到的汽车价款才等于他对汽车的估值。仔细观察若 $\overline{p} \geq 15\,000$，则有低品质汽车的卖方也认为卖掉汽车是最佳策略。如果在汽车是高品质的条件下，$\overline{p} \geq 20\,000$，那么选择惜售是最佳策略。总的来说，当 $15\,000 \leq \overline{p} \leq 20\,000$ 时，卖方的策略是最优的。

至于买方的信念，如果她看到汽车在市场的售价是 $\overline{p}$，那么根据卖方的策略，被出售的汽车一定是低品质的。一致性要求给高品质汽车赋予零概率，并且运用贝叶斯法则，我们可以求出给汽车是低品质的赋予的后验概率是 0.375，即 0.3/(0.3+0.5)，中等品质的后验概率是 0.625，即 0.5/(0.3+0.5)。考虑到这些信念，当且仅当：

$0.625 \times (18\,000 - \overline{p}) + 0.375 \times (12\,000 - \overline{p}) \geq 0$，或 $\overline{p} \leq 15\,750$

时，买方以价格 $\overline{p}$ 购买的策略才是最佳策略②。

总之，买方只有在弄清汽车是低品质或中等品质的情况下，才愿意以 15 750 的价格购买，而当汽车是中等品质并且价格至少是 15 000 时，卖方才愿意出售汽车。这样，若 $\overline{p}$ 的取值在 15 000 和 15 750 之间（如 15 499），则上述的策略组合和信念集合构成了一个精炼贝叶斯纳什均衡。到此我们终于解决了这个问题！

上述分析使我们对非对称信息是如何影响某一市场的高效运作有了更为深刻的认识。在对称信息下，向买方出售高品质的汽车可以使交易的双方共赢。例如，买方以 21 000 的价格购买的汽车对她来说可以产生 24 000 的价值（这样她就可以得到 3 000 的收益），而卖方以这

① 从专业角度来说，我们也应该明确在汽车不被出售的条件下，买方对汽车质量的信念。注意一致性暗示着这些信念认为汽车是高品质的可能性为1，因为根据卖方的策略，只有高品质汽车的卖方才会惜售。可是，在汽车不被出售时，买方无法采取任何行动，所以描述这些信念就没有什么意义。——作者注

② 出于完整性的考虑，我们也必须分析当价格不同于 $\overline{p}$ 时买方的策略是最优的。这一情形也很容易证明。——作者注

样的价格出售也得到比他对汽车的估值高出 1 000 的收益。遗憾的是，出售高品质汽车的情形不是一个均衡。这些高品质汽车之所以退出市场是因为卖方不能以一个较高的价格销售这些汽车。既然市场上的所有汽车都必须以相同的价格出售，那么低品质的汽车会减少买方愿意支付的价格，但这一价格又使得有高品质汽车的卖方不愿意出售。这样，买方和卖方之间信息的不对称使得劣等车将高品质汽车驱逐出市场。

11.3 理解测验

在只出售低品质汽车的情形下，求出一个精炼贝叶斯纳什均衡。

恋爱

人类社会中最重要的一种关系就是恋爱，在这个过程中人们为了更长久地在一起，彼此熟悉，甚至产生亲密关系。

从社会学和生物学的角度来看，恋爱当中最重要的就是性关系的作用和出现的时机。是在订婚之前发生性关系，还是在订婚后结婚前？亦或者只有在结婚后才发生性关系？历史上，社会和伦理规范允许在恋爱期间发生性行为。在 20 世纪中叶，美国颁布的条例和规范大都不承认婚前性行为的合法性。事实上，在 20 世纪 40 年代发布《金赛性学报告》（*Kinsey Reports*，这是一项富有开拓性意义的报告）之前，美国民众普遍认为只有婚后才能发生性行为。然而，这一报告却指出美国有近一半的情侣会在婚前发生性行为。

在一个女性必须坚守婚前贞操的时代，一个女人必须为因冲动发生的婚前性行为付出一定的代价。婚前性关系相当普遍的现象表明通过订婚来实现婚约降低了女性的风险。但是如何防止一个男人为了发生性关系而许诺结婚，之后又不兑现呢？如果这个男人真的愿意结婚又如何呢？他该如何表达自己的真挚感情以让人相信呢？或许可行的办法只能是让自己的表达方式无法被虚情假意的男人所模仿。下面，我们用博弈理论来探究上述种种问题和其他严肃的社会问题。

这一情形中的私人信息是这个男人（假定是杰克）是否很在乎这个女人（假定是罗斯），并且愿意娶她；类似的，罗斯是否十分在乎杰克，并且愿意嫁给他。只有一个人知道他或她是否真的爱对方时，这种爱出现的概率是 p，其中 $0 < p < 1$。这样，他们彼此心仪对方的概率是 $p \times p$，或 p^2，即杰克爱罗斯而罗斯也爱杰克的概率。注意我们正进一步加大信号传递博弈的难度，因为本例中的接收者罗斯也有私人信息。我们探索的方法会适用于发出者和接收者双方。①

杰克和罗斯的决策次序如下：杰克先决定是否提出亲密要求。然而，由于言语往往缺乏说服力，所以我们假定杰克在表明心意的同时也向对方送了一件礼物。为了回应对方送出的礼物，罗斯可以选择接受或拒绝，如果接受就表明愿意亲密接触。这一连串事情结束后，他们要么结婚（如果彼此爱着对方），要么不结婚（如果有一方或双方都不爱对方）。假定结婚决定（不予明确表示，相反会暗含在收益中）与是否发生性行为无关。

表 11—4 表述的是所有可能结果的收益。杰克的收益取决于他的礼物是否被接受（然后他们发生性关系）和他们是否钟爱对方（然后结婚）。不管是否爱罗斯，杰克都想和她发生关系。他从发生关系中获得的收益用 $s\,(s>0)$ 表示，而他买礼物的损失用 $c\,(c>0)$ 表示（只有罗斯接受礼物时才会发生这一情况）。可是，如果他最终娶了她，那么他也会从对方接受礼物中获得一些快乐，所以在这种情况下他的净损失只是 $\frac{c}{2}$。若杰克和罗斯彼此相爱并且结婚，则他赋予结婚的价值是 $m\ (m>0)$。总之，如果杰克和罗斯发生关系并且他们结婚（因为事实证明他们彼此相爱），那么他的收益是 $m+s-\frac{c}{2}$。如果他和对方发生关系，但并没有和对方结婚，那么他的收益是 $s-c$。如果他们结婚而没有发生关系，那么他的收益是 m。最后，如果他既没有发生关系也没有结婚，那么他的收益为零。需要注意的是只有当他们彼此相爱时才会结婚，而这一点由命运决定而不是由选择决定。

① 这一博弈由三个阶段组成：（1）“自然”决定发出者和接收者的类型。（2）发出者根据自己的类型做出行动。（3）接收者根据自己的类型，观察发出者的行动，通过获取的新信息更新她的信念，然后采取行动。发出者的策略是对每个可行类型采取行动，而接收者的策略是每个由发出者类型和接收者类型组成的可行类型集合采取行动。——作者注

表 11—4　恋爱博弈中的收益

是否送礼物和发生关系	杰克是否爱罗斯	罗斯是否爱杰克	杰克的收益	罗斯的收益
是	是	是	$m+s-\frac{c}{2}$	$m+s+v$
是	是	不是	$s-c$	$v-u$
是	不是	是	$s-c$	$v-u$
是	不是	不是	$s-c$	$v-u$
不是	是	是	m	m
不是	是	不是	0	0
不是	不是	是	0	0
不是	不是	不是	0	0

和杰克一样，罗斯赋予他们彼此相爱和结婚的价值是 m，但是她不想和将来不会结婚的人发生关系。罗斯从选择接受礼物，与杰克发生关系并和他结婚得到的收益是 $m+s+v$，其中 $v\,(v>0)$ 是礼物的价值。然而，如果她和杰克发生关系但并没有和他结婚（出现这一状况或者是因为她不爱对方或者是对方不爱她），那么她的损失是 $v-u$，其中 $u>0$，这一点可以被理解为失去贞洁的损失。她既不发生关系也不结婚的收益是零。

注意杰克和罗斯在性关系和礼物方面的偏好不同。对于杰克，买礼物是一种损失，而对罗斯而言，接受礼物是一种收益。在性关系方面，无论他和罗斯是否相爱，杰克都想发生关系，而罗斯只有在要和对方结婚时才会愿意发生关系（这种情况也只发生在他们彼此相爱的情况下）。

图 11—3 描述的就是贝叶斯博弈的情形。“自然”先决定杰克是否爱罗斯和罗斯是否爱杰克。正如他们的信息集合表示的那样，杰克知道在他决定是否发生性关系时自己是否爱罗斯，但他不知道罗斯的态度会如何。如果杰克提出结婚，尽管罗斯知道自己对杰克的感情，但当她必须决定是否和他发生关系时，并不知道杰克的真实意图（他真的爱我吗？）。

我们接着分析下面的策略组合和信念集合：

■ 杰克的策略：

如果我爱罗斯，就送给她一个礼物。

如果我不爱罗斯，就不送给她礼物。

■ 罗斯的策略：

如果我爱杰克并且他送给我一个礼物，那么接受这个礼物。

如果我不爱杰克而他送给我一个礼物，那么拒绝这个礼物。

■ 罗斯的信念：

若杰克送礼物，那么他爱我的概率为 1。

若杰克不送礼物，那么他不爱我的概率为 1。

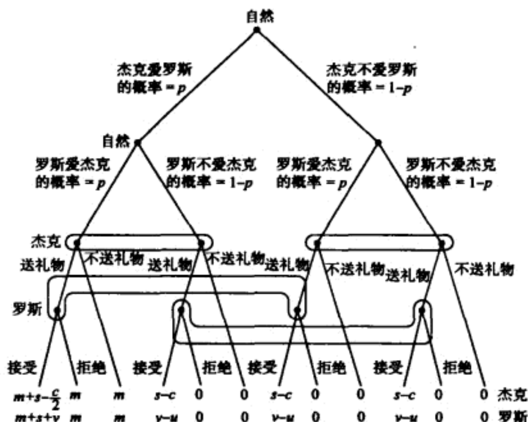


图 11—3　恋爱博弈

在证明上述情形是否是精炼贝叶斯纳什均衡之前，先分析杰克的策略。他不爱罗斯的话，那么他的策略不允许他送礼物，当：

$$0 \geqslant p\times(s-c)+(1-p)\times 0 \text{ 或 } c \geqslant s \tag{11.2}$$

时，选择不送礼物是一个最佳策略。而不送礼物的收益是零，因为在这种情况下他不和她发生关系，而且他也不会娶她（因为他不爱她）。倘若他送礼物，则根据出罗斯爱他的概率是 p，这样她会接受礼物（根据她的策略），尽管他们不会结婚（因为杰克不爱她）。由此他得到的收益是 $s-c$，即发生关系的价值减去买礼物的损失。根据概率 $1-p$，罗斯会拒绝礼物因为她不爱他，这种情况下他的收益为零。因而，杰克不爱罗斯时选择不送礼物是他的最佳策略，特别是当礼物很贵重时，即 $c \geqslant s$。

现在假定杰克爱罗斯。他的策略要求他送礼物，这一策略当：

$$p\times\left(m+s-\frac{c}{2}\right)+(1-p)\times 0 \geqslant p\times m+(1-p)\times 0 \text{ 或 } 2s \geqslant c \tag{11.3}$$

时，是最佳策略。不等式的左边是选择送礼的期待收益，右边是不送礼物的期待收益。按照罗斯的策略，只有当她爱杰克时才会接受礼物。这样，如果送礼物，那么被接受的概率是 p，并且杰克的收益是 $m+s-\frac{c}{2}$，

因为他从发生关系中得到 s，从和罗斯结婚得到 m（他们相爱）和送礼物净损失 $\frac{c}{2}$。在概率是 $1-p$ 的情况下，罗斯不爱他，而且因为她拒绝礼物所以不会和他结婚，那么杰克的收益就是零。如果他没有送礼物，在罗斯爱他的前提下，他仍然会娶她（这一情况发生的概率是 p），否则收益为零。因而，如果杰克爱罗斯，那么当礼物不太贵（$c \leqslant 2s$）时，送给她礼物是最佳策略。

将公式（11.2）和公式（11.3）合并，我们推出精炼贝叶斯纳什均衡出现的条件是 $2s \geqslant c \geqslant s$。如果礼物太贵，那么杰克不愿意仅仅为了发生关系而送礼物。如果礼物非常便宜，那么他很愿意给他未来的新娘送礼物，从而可以在婚前发生性关系。

现在分析罗斯的策略。首先，注意她的先验信念和后验信念是一致的。给定杰克的策略，她可以从他的行动中准确推出他的情感，即她认为只有杰克爱她时才会送礼物。因而，如果罗斯爱杰克，并得到他的礼物，那么当 $m+s+v \geqslant m$ 或 $s+v \geqslant 0$ 时，选择接受是最佳策略，这一表达式是成立的，因为 s 和 v 都是正数。送礼物表明他爱她，如果她也爱他，那么她选择接受礼物和发生关系（确认他们会结婚）得到的收益是 $m+s+v$。如果她拒绝礼物，但仍希望嫁给他，那么她的收益是 m。这样她会选择接受礼物。

假定罗斯不爱杰克。她的策略要求她拒绝礼物，那么当 $0 \geqslant v-u$ 或 $u \geqslant v$ 时，这一策略的确是最佳策略，由此得到的收益为零。选择接受礼物，她的收益是 $v-u$，因为虽然得到礼物的价值 v，但因为成为失去贞洁的单身女人会遭受损失 u。由于拒绝杰克是最佳策略，我们推出 $v \leqslant u$。在罗斯看来这一礼物越贵越好，如果很贵的话，就算她不打算嫁给杰克也会接受礼物。

总之，只有当杰克爱罗斯并且送给她礼物，以及只有当罗斯爱杰克并且接受礼物的情形才是精炼贝叶斯纳什均衡。当（1）杰克不愿意仅仅为了发生关系而买礼物（$c \geqslant s$），（2）在礼物不太贵的前提下，如果杰克想让她成为自己的新娘，那么他愿意送礼物，（3）罗斯认为礼物的价值不足以让她在不期待结婚的条件下接受礼物时，也是一个精炼贝叶斯纳什均衡。因而，礼物应该贵，但不能太贵，而且也应该奢华（这样她就不认为礼物很贵）。

我们分析一下这些条件什么时候成立。首先，注意无论是否爱罗斯，杰克都想和她发生关系，这样他的收益会因为发生关系而增加 s。两个类型的杰克（真诚的和虚假的）之间的区别是他发生关系的损失。对于没有爱的（或虚假的）杰克，他的损失是礼物的价格 c，而对于充满深情的（或真诚的）杰克，他的损失是 $\frac{c}{2}$，因而较小；因为他知道只有罗斯也爱他，她才会接受礼物，而且他也从给自己未来的新娘送礼物中获得利益。图11—4表示的是 c 取不同值的时候，虚假的杰克买礼物的损失 c 和真诚的杰克买礼物的损失 $\frac{c}{2}$。此外，图中也描述了发生关系的价值 s。

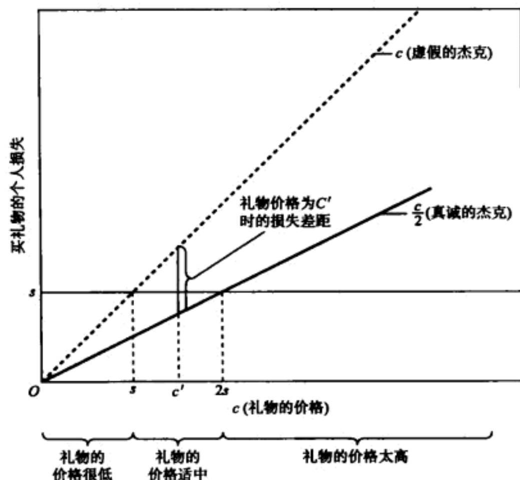


图 11—4 用礼物表达爱意的损失

虚假的杰克和真诚的杰克买礼物产生的损失差距是 c 线和 $\frac{c}{2}$ 线之间的垂直距离。例如，当礼物的价格是 c' 时，图11—4中描述出了这个差距。注意损失的差距随着礼物的标价而增加。礼物价格不太高时（比如 $c<s$），没有分离均衡。问题是无论是真诚的还是虚假的，杰克都愿意在礼物能够诱使罗斯和他发生关系的条件下送给她礼物。然而，虚假的杰克送的礼物太过廉价而无法让人相信他爱的表达。不过，如果礼物较贵（但不太贵），那么虚假的杰克和真诚的杰克买礼物产生损失的差距就很大，这时，只有真诚的杰克愿意买礼物，并且当 $s \leqslant c \leqslant 2s$ 时会发生这种情况。①

我们从中得出的启示是如何满足出现分离均衡的条件。重要的不是礼物价格昂贵，而是与杰克爱罗斯时相

① 如果礼物太贵（$c>2s$）的话，那么任何一种类型的杰克都不会买礼物，这样就又不存在分离均衡。——作者注

比，杰克认为不爱罗斯时送礼物更贵。也就是说，不同类型的杰克之间差别化的损失才是关键点。将这一博弈和二手车博弈（当中无论卖方的类型是什么，标出更高价格的利益是相同的）做一比较。在后一种博弈中，卖方标价对他收益的影响不会随着汽车的质量而变化，所以拥有低品质汽车的卖方会倾向于和拥有高品质汽车的卖方一样标出相同的高价。正是由于很容易出现模仿，所以分离均衡不存在。



点拨

某一类型的发送者可以通过选择一个对其他类型的发送者来说相对成本更高的行动——这样他们就不会选择那一行动，从而有效地将自己和其他类型的发送者区别开来。因而，成功地表示出某一行动的损失会随着发送者类型的变化而变化是很重要的。

在美国的许多州，如果一名妇女因为虚假婚约而给自己带来损失的事实得到证实，那么她可以按照有关“毁弃婚约”的法律控告悔婚的未婚夫。可以提起诉讼的威慑力很可能防止了一些悔婚现象的发生。然而，这些法律从 20 世纪 30 年代开始被废除，结果被抛弃的妇女无法控告自己的未婚夫。一个相当有趣的现象是就在这同一时期，用钻石戒指求婚的习俗出现了。钻石戒指也许就是我们所讲模型中的昂贵而又奢华的礼物，这一礼物使得发生婚前性行为的现象开始出现，这一点在《金赛性学报告》中得到了证实。（用宝石制作订婚戒指而不是结婚戒指的现象的确很有趣。）一个打算结婚的男人可能愿意给将成为他未来妻子的女人准备一个钻石戒指，但是一个只对虚假婚约感兴趣的男人会认为准备一个钻石戒指的代价过于昂贵。

边缘政策博弈

20 世纪 50 年代，美国国务卿约翰•福斯特•杜勒斯（John Foster Dulles）提出不怕走到战争边缘，但要学会走到战争边缘，又不卷入战争是边缘政策的必要艺术。美国当前正努力控制朝鲜实施研发核武器计划的局面，正和 1956 年的情形一样需要运用这种必要艺术。美国担心的是朝鲜可能会针对韩国或其他一些国家使用核武器。虽然任何一个理智的领导人都不会那么做（因为这样做会导致来自实力更强大国家的大规模报复），但这确实是个问题：金正日是理智的吗？面对这样的不确定性美国该如何采取行动？尽管我们不能提供一个供美国总统采纳的解决办法，但我们可以从博弈论的角度来看待这一问题，从而明确一些相关的注意事项。①

这个博弈的参与者是乔治•布什和金正日。他们的决策次序如图 11—5 所示。“自然”先决定金正日是理智的还是疯狂的，其中他疯狂的概率是 0.25。了解自己的心理状态之后，金正日决定是坚定研发核武器的立场还是对美国和联合国的要求做出让步。如果他做出让步，那么危机（博弈）结束。如果他坚定立场，那么美国必须决定是坚定立场还是做出让步。如果美国做出让步，那么博弈结束。但如果美国选择坚定立场，那么朝鲜必须决定是否采取敌对行动（如使用核武器）。

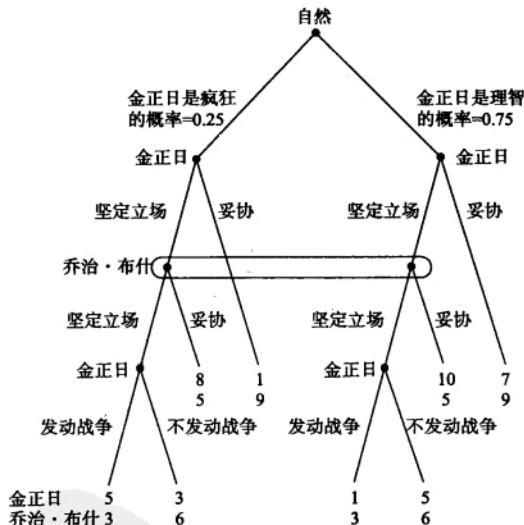


图 11—5　乔治·布什和金正日之间进行的边缘政策博弈

如前所述，金正日有两种可行类型：理智型和疯狂型。若是理智的，则他（按照最好到最差的顺序）排列出以下四种可能的结果：他坚定立场，布什让步；他妥协；他和布什都坚定立场时他不卷入战争；他和布什都坚定立场时他参加战争。相反，如果金正日是疯狂的，

① 分析这个例子需要掌握附录中所描述的贝叶斯法则的一般形式。——作者注

那么他的偏好排序为：他坚定立场，布什让步；他和布什都坚定立场时他参加战争；他和布什都坚定立场时他不卷入战争；他让步。这两种类型最大的不同点是若对手不让步那么疯狂的金正日就愿意参战。

在描述精炼贝叶斯纳什均衡时，允许对两位领导人的顺序做任意排列。因为布什只有一个信息集合，他的（混合）策略是选择坚定立场的概率。对于金正日而言，他有四个信息集合。当他面对第一个信息集合（坚定立场或让步）时，选择的策略会告诉他采取什么行动。当他坚定立场而美国也不让步（参战还是不参战）时，采取的行动取决于他的类型（理智的还是疯狂的）。

分析下面的一个策略组合：

■ 金正日的策略：
若是疯狂的，则选择坚定立场；若朝鲜和美国都选择坚定立场，则选择参战。
若是理智的，则选择坚定立场的概率为 k；若朝鲜和美国都选择坚定立场，则选择不参战。
■ 乔治·布什的策略：选择坚定立场的概率为 b。

既然信念是表述精炼贝叶斯纳什均衡的组成部分，那么先描述与金正日的策略相一致的布什的信念。布什无法推测金正日的类型，因为虽然金正日疯狂时会选择坚定立场，但是他理智时选择坚定立场的概率为 k。在这种情形下，我们需要运用贝叶斯法则来推断出布什的信念是什么。根据这一法则，我们可以得出：

$$\text{概率（根据他采取坚定立场推断是疯狂类型）}$$
$$=\frac{\text{概率（疯狂）}\times\text{概率（根据他是疯狂采取坚定立场）}}{\text{概率(疯狂)}\times\text{概率(基于他是疯狂型采取坚定立场)}+\text{概率(理智)}\times\text{概率(基于他是理性采取坚定立场)}}$$
$$=\frac{0.25\times1}{0.25\times1+0.75\times k}=\frac{0.25}{0.25+0.75k}$$

若理智的金正日不选择坚定立场（即 k=0），则：

$$\frac{0.25}{0.25+0.75\times0}=1$$

这样布什可以从金正日采取的坚定立场准确推断出金正日是疯狂的。随着 k 的增加，更有可能是理智领导选择坚定立场，那么基于他采取的坚定立场，金正日是理智型的概率增大。当 k=1 时，这种概率达到最大值 0.75。出现这种情况绝非偶然，这一数值等于先验概率，因为倘若无论金正日是什么类型都肯定选择坚定立场，那么他采取坚定立场的行动无法为判断他的心理状态提供任何信息。这样，贝叶斯法则使得后验信念和先验信念相一致。

相反，若布什观察到金正日让步，则他的后验信念认为这一领导人是理智的概率为 1，因为一个疯狂的领导不会让步。这一结果也可以从贝叶斯法则得出：

$$\text{概率（根据他采取让步推断是疯狂类型）}$$
$$=\frac{\text{概率（疯狂）}\times\text{概率（根据他是疯狂型采取让步）}}{\text{概率(疯狂)}\times\text{概率(基于他是疯狂型采取让步)}+\text{概率(理智)}\times\text{概率(基于他是理性采取让步)}}$$
$$=\frac{0.25\times0}{0.25\times0+0.75\times(1-k)}=0$$

将这些一致的信念和下面的策略组合，即：

■ 乔治·布什的信念：
若金正日选择坚定立场，则他疯狂的概率为 $\frac{0.25}{0.25+0.75k}$。
若金正日选择让步，则他疯狂的概率为零。

现在我们求出所给策略组合中 b 和 k 的值，从而使这一策略组合和之前的信念集合成为一个精炼贝叶斯纳什均衡。假定朝鲜选择坚定立场，那么当且仅当：

$$5=\left(\frac{0.25}{0.25+0.75k}\right)\times3+\left(\frac{0.75k}{0.25+0.75k}\right)\times6 \qquad (11.4)$$

时，布什采取何种行动都可以。无论金正日是什么类型，布什选择让步得到的收益是 5；金正日是疯狂类型（战争随后而至）时，布什选择坚定立场得到的收益是 3；金正日是理智的时候，布什选择坚定立场得到的收益是 6。通过求公式(11.4)中 k 的值，我们得到 k=0.67。因而，如果当金正日是理智的时候选择坚定立场的概率为 0.67，那么无论布什选择让步还是坚定立场得到的期待收益都相同。

最后，分析一下金正日的策略。很容易证明当他是疯狂类型时，所选的策略是最佳策略。如果选择让步，那么他得到的收益最少，很显然他会因此选择坚定立场。在布什也选择坚定立场的情形下，他选择参战得到的收益 5 要高于不参战得到的收益 3。

现在假定金正日是理智的。假设他必须决定和布什都选择坚定立场时是否参战，那么和参战得到收益 1 的情形相比，选择不参战并得到收益 5 的策略的确是最佳策略。最后需要分析的信息集合是金正日在坚定立场和

让步之间做出选择。回顾之前我们假定面对这一信息集合时，他采取的行动不分先后。（特别是，我们已经证明布什的行动不分先后时他选择坚定立场的概率为 0.67。）这一策略要成为最佳策略，必须满足：

$$7 = b \times 5 + (1-b) \times 10 \qquad (11.5)$$

等式的左边是金正日选择让步得到的收益，右边是选择坚定立场得到的期待收益（由布什选择坚定立场还是让步的可能性决定）。求公式（11.5）中 b 的值，我们得到 b=0.6。均衡要求布什选择坚定立场的概率是 60%。

总之，均衡做出的预测是如果金正日是疯狂的，那么他肯定选择坚定立场；如果他是理智的，那么他选择坚定立场的概率是 67%。面对朝鲜的坚定立场，布什选择坚定立场的概率是 60%。而出现僵局时（即哪一国都不选择让步），只有一个疯狂的金正日会发动战争。

本章总结与练习题

本章总结

本章重点评析了具有两个特点的冲突问题，即参与者有私人信息和按序列行动。序贯行动的重要性在于第一个采取行动的参与者会泄露她知道而其他参与者不知道的信息。第二个行动的参与者可以从她观察到的行动收集有关参与者的类型，并且根据这一信息确定要采取的行动。但如果这一情况是连续性的，那么先采取行动的参与者可能会改变行动从而误导那些从她的行动获取信息的参与者。本章的目标是分拣出这些各异的因素，明确一些解决方案，这些方案中所有的参与者都以自己利益最大化为前提而采取行动，而且没有参与者被误导。虽然这并不意味着一个参与者一定可以弄清另一个参与者知道的信息，但至少不会轻易相信虚假的东西。

我们之所以在信号传递博弈的情境下着手进行这一分析是因为该博弈是包含所有这些特征的最简单的结构。信号传递博弈中有两个参与者，一个是有一些私人信息并首先采取行动的发送者，另一个是观察发送者行动的接收者，他通过更新对发送者私人信息的信念来确定自己采取的行动。使用的解决方案是**精炼贝叶斯纳什均衡**。大体而言，如果每个参与者（1）根据她的信念（**序贯理性**）采取最佳行动，（2）基于发送者行动而形成的接收者的后验信念能够正确描述发送者采取的最佳行动（**一致信念**），那么一个策略集合和一个接收者对发送者私人信息的形成的后验集合就构成了精炼贝叶斯纳什均衡。换句话说，当接收者更新自己的信念以回应所观察到的发送者的行动时，就表明他识别出发送者企图误导他的动机。

出现一个均衡要满足的条件是发送者的行动（1）完全透露出她知道的所有信息（**分离均衡**），（2）透露一部分她知道的信息（**半分离均衡，或半混同均衡**），或（3）不透露一点信息（**混同均衡**）。在恋爱博弈中，一名求婚者用贵重的礼物求婚的行动向他的未婚妻透露出他真心爱她的信息，以及在管理培训生博弈中，实习生个人努力程度透露出她是否勤于工作时的博弈情形都是分离均衡。而半分离均衡则出现在二手车博弈，其中卖方只出售低品质或中等品质汽车。由于卖方定价与汽车质量无关，因而虽然买方不知道汽车是低品质的还是中等品质的，但清楚汽车不是高品质的，因为拥有高品质汽车的卖方不会选择在市场上出售。边缘政策博弈中也透露出了部分信息。面对美国的威胁，金正日选择坚定立场的行动虽然没有透露他是疯狂的还是理智的，但是布什总统可以据此更新自己的信念，推断出金正日是疯狂类型的概率较大。

从对**信号传递博弈**的分析中可以得到什么启示呢？第一个启示是发送者提供信息的行动取决于发送者的类型是如何影响他的行动偏好的情形是否是一个均衡。无论发送者的类型是什么，如果他的行动得到的收益都相同，那么就不太可能出现分离均衡。例如，高品质汽车的卖方和低品质汽车的卖方以较高的价格出售时得到的利益相同的话，那么在这种情形下，卖方的收益就等于汽车的销售价。所以，低品质汽车的

卖方就会倾向于模仿高品质汽车卖方的定价。因而，我们看出二手车的标价是和它的质量无关的，那么也就无法从价格辨别汽车的真实状况。相反，在恋爱博弈中，如果一个男人爱一个女人并期望娶她，那么他送给她一个贵重的礼物的损失就较小。在这一情形下，礼物就可以真实地表示出他对对方的爱。

第二个启示是一个参与者要想成功地表明自己的类型，那么她可能需要采用极端点的行为方式使得其他类型的参与者无法模仿。比如，在管理培训生博弈中，为了和懒散的实习生区别开来，勤勉的实习生必须比平常更努力才行。勤勉型的实习生不得不使自己努力的程度达到懒散型实习生选择不愿模仿的程度，即使达到这一程度能使经理误认为懒散的实习生是勤勉的。

练习题

1. 回到 11.2 节所讲的次品博弈，分析类似的情形：汽车质量的类型有两种，即低品质和高品质。买方起初认为汽车属于高品质类型的概率为 q，低品质的概率为 $1-q$。如果卖方售车，那么她的收益是买方支付的汽车价款。如果不出售汽车，那么假定卖方对高品质汽车的估值是 10 000 美元，低品质汽车的估值是 6 000 美元，这些估值就是她得到的收益。如果买方对高品质汽车的估值是 12 000 美元，低品质汽车的估值是 7 000 美元，那么他的收益就是汽车的价值（即 12 000 或 7 000）减去支付的价格；若他不买车，则收益为零。求出使精炼贝叶斯纳什混同均衡成立的 q 的值。

2. 求出 11.2 节所讲的恋爱博弈中存在的一个精炼贝叶斯纳什混同均衡。

3. 一个零售商出售一种产品的质量有低和高两种类型，而产品质量根据产品正常使用的概率来划分。假设高品质产品正常使用的概率是 h，低品质产品正常使用的概率是 l，并且 $0<l<h<1$。制造商的类型就是其产品质量（低品质或高品质），其中消费者认为产品是高品质的概率为 r，且 $0<r<1$。零售商销售两种质量类型产品的成本是 c（只有卖掉产品才会有成本）。“自然”决定零售商产品的质量类型后，零售商必须同时决定产品的价格和是否承诺退款。零售商的定价可以是任何一个非负数。承诺退款是指如果产品无法使用（高品质产品出现的可能性是 $1-h$，低品质产品出现的可能性是 $1-l$），那么零售商必须退还消费者所购买产品的价款。零售商的收益是他的期待利润。所以，如果他的产品定价是 p，并且不承诺退款，那么他的收益是 $p-c$；如果承诺退款，那么他出售高品质产品得到的期待收益是 $h\times p-c$，出售低品质产品得到的收益是 $l\times p-c$。看到产品标价和是否有退款承诺后，一个消费决定是否购买。在表述消费者的收益时，先注意若产品能正常使用，则消费者认为产品的价值是 $v\ (v>0)$，而产品不能正常使用的话，那么这个产品就没有价值。如果她的购买价格是 p，那么她的收益是（1）$v-p$（产品能正常使用时），（2）$-p$（产品不能正常使用，并且没有退款承诺时），（3）0（产品不能正常使用，有退款承诺时）。消费者不买产品的收益是零。最后，假定 $0<c<l\times v$。求出一个精炼贝叶斯纳什均衡，其中零售商选择只对高品质产品承诺退款，消费者不管有没有退款承诺都购买高品质产品。（提示：零售商不承诺退款时可以以较低的价格销售产品。）

4. 乔治·布什总统和伊拉克领导人侯赛因·萨达姆处于策略对峙中。侯赛因的类型决定是否有大规模杀伤性武器，其中，他有大规模杀伤性武器的概率是 w，并且 $0<w<\frac{3}{5}$。了解自己的类型后，萨达姆决定是否允许武器核查。如果他允许，那么假定核查人员发现大规模杀伤武器会上报，没有发现大规模杀伤性武器则不上报。如果他不允许核查，那么无法判定他是否有大规模杀伤性武器。这种情况下，布什总统决定是否入侵伊拉克。图 11—6 表述的就是这一博弈的拓展形式。注意如果侯赛因允许武器核查，那么布什就可以弄清侯赛因的类型，否则将无法判断。

侯赛因的收益是无论他是否有大规模杀伤性武器，他对可能出现结果的排序如下（从最好到最差的顺序）：不允许核查，不出现入侵；允许核查，不出现入侵；不

允许核查，出现入侵；允许核查，出现入侵。这样，侯赛因会选择不允许核查，但也最不希望出现入侵。布什的偏好排序很大程度上取决于侯赛因是否有大规模杀伤性武器。如何侯赛因有大规模杀伤性武器，那么布什选择入侵；如果他没有大规模杀伤性武器，那么布什不选择入侵。写出布什的一致信念并求出 $b(0<b<1)$ 和 h（$0<h<1$）的值，使得下面的策略集合成为一个精炼贝叶斯纳什均衡。

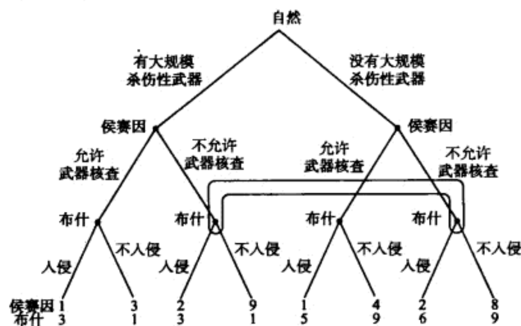


图 11—6 大规模杀伤性武器博弈

■ 侯赛因的策略：

如果我有大规模杀伤性武器，那么不允许武器核查。

如果我没有大规模杀伤性武器，那么允许核查的可能性是 h。

■ 布什的策略：

如果侯赛因允许核查，并发现大规模杀伤性武器，那么选择入侵。

如果侯赛因允许核查，没有发现大规模杀伤性武器，那么选择不入侵。

如果侯赛因不允许核查，那么入侵的可能性是 b。

5. 我们现在证明一个大学学位是如何帮你找到一个更好的工作的，即使它不能使你成为一个更优秀的员工。分析发生在一个未来员工（我们称之为求职者）和正式员工之间的一个两人博弈。求职者的类型是她的智商类型，即低、中等和高三种，三种类型的概率分别为$\frac{1}{3}$，$\frac{1}{2}$和$\frac{1}{6}$。求职者了解她的类型后，决定是否上大学。如果求职者不那么聪明，那么她获得大学学位的代价就更大，因为如果她想毕业就必须更加努力。假定一个高、中等或低智商的求职者获得大学学位的代价分别是 2,4 和 6。雇主决定是否给求职者提供一个经理或职员的职位。求职者获得经理一职的收益是 15，而得到职员一职的收益是 10。这些收益和求职者的类型无关。雇主聘用某人为职员得到的收益是 7（无论这个人是否聪明以及是否有大学学位）。如果求职者被聘为经理，那么雇主的收益就会随着求职者的智商从 4 增加至 6 或 10（这一点取决于求职者的智商是低、中等还是高）。注意雇主的收益和求职者是否有大学学位无关。图 11—7 表述的就是这一博弈的扩展形式。证明低智商的学生不上大学与中等和高智商学生上大学的情形是精炼贝叶斯纳什均衡。

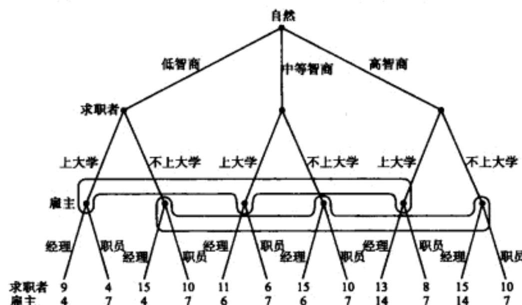


图 11—7 "上大学"信号传递博弈

6. 一家新开张的餐厅老板打算打广告以吸引更多的消费者就餐。在这个贝叶斯博弈中，"自然"决定餐厅的品质只有高低两种类型。假定每个品质出现的概率相同。老板了解品质类型之后，决定打广告的费用。设 A 为广告开销。为了便于分析，假定只有一个顾客。顾客看到餐厅的广告宣传之后，更新自己对餐厅品质的信念，然后决定是否去这家餐厅就餐。（可想而知，观察到的 A 包括在地方电视台有多少条插播广告，在报纸和广告栏上登载了多少条广告。）假设一顿饭的价格是 50 美元，而消费者认为高品质饭菜的价值是 85 美元，低品质饭菜的价值是 30 美元。去这家餐馆就餐的消费者如果发现饭菜的品质低，所以得到的收益是−20（＝30−50）；如果发现饭菜品质高，那么得到的收益是 35(=80−50)。而且，了解了餐厅的餐饮水准后，消费者会选择再次光顾。因而，消费者去高品质餐厅就餐得到的收益实际上是 70(=2×35)。对于餐厅老板来说，假定提供一顿饭的成本是 35（无论是低品质还是高品质）。如果餐厅品质高，消费者去餐厅就餐，以及餐厅的广告开销是 A，那么得到的收益是 $20\times(50-35)-A=30-A$。如果餐厅品质低，消费者去就餐，以及餐厅的广告开支是 A，那么得到的收益是 $(50-35)-A=15-A$。表 11—5 所列的就是这些收益。如果消费者不去就餐，那么她的收益为零，餐厅老板的收益为$-A$。

表 11—5　餐厅博弈：一个消费者去餐厅就餐得到的收益

餐厅的品质	餐厅老板的收益	消费者的收益
低	15−A	−20
高	30−A	70

a. 写出一个分离精炼贝叶斯纳什均衡。

b. 在一个分离精炼贝叶斯纳什均衡中，餐厅打广告的最高费用是多少？最低费用又是多少？

c. 写出一个混同精炼贝叶斯纳什均衡。

d. 在一个混同精炼贝叶斯纳什均衡中，餐厅打广告的最高费用是多少？

7. 分析图 11—8 所示的信号传递博弈情形。“自然”选择三种类型发送者中的一种，发送者在了解自己的类型后，采取三种行动中的一种。接收者观察发送者的行动，但不是她的类型，然后选择两种行动中的一种。推出一个半分离精炼贝叶斯纳什均衡，其中发送者的类型是 t_1 或 t_2 时选择相同的行动，当她的类型是 t_3 时选择不同的行动。

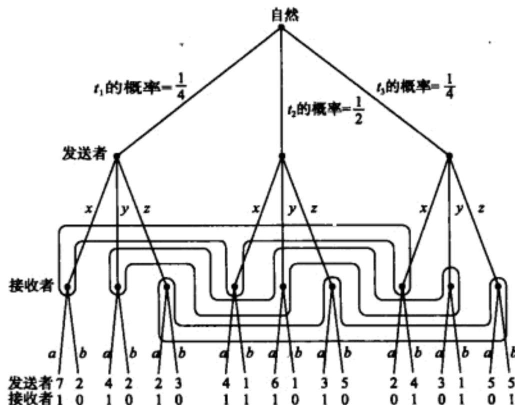


图 11—8

8. 解出图 11—9 所示信号传递博弈中的所有分离精炼贝叶斯纳什均衡。

9. 分析图 11—10 所示的信号传递博弈情形。

a. 推出所有的分离精炼贝叶斯纳什均衡。

b. 推出所有的混同精炼贝叶斯纳什均衡。

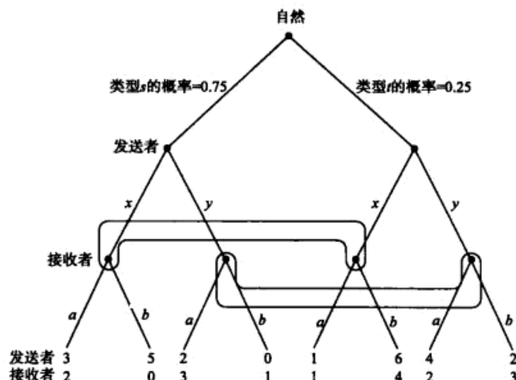


图 11—9

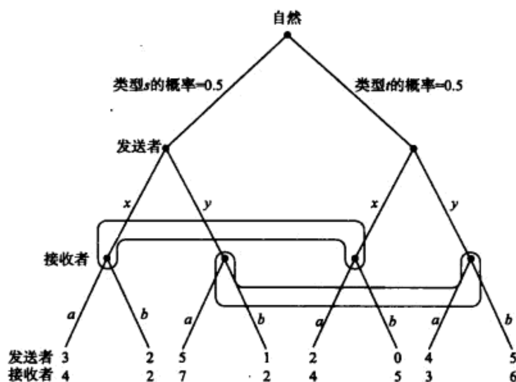


图 11—10

附录

贝叶斯法则和信念更新

思考一个大学有 4 000 名学生，其中女生有 2 100 名，男生有 1 900 名。表 11—6 根据性别和专业领域将全体学生分为四个类别：未公开、人文学科、社会科学和自然科学。这些分类并不相互交叉，也就是说，一名学生只属于一个类别，不能多也不能少。

表 11—6　某大学学生人数

	未公开	人文科学	社会科学	自然科学	总计
男性	250	350	600	700	1 900
女性	150	450	900	600	2 100
总计	400	800	1 500	1 300	4 000

这一表格可以用来对随机选取学生的特点赋予不同的概率。例如，随机选取的学生是女生的概率为 0.525 $\left(=\frac{2\,100}{4\,000}\right)$，因为在 4 000 个学生中女生有 2 100 人。类似的，随机选取学生的专业是社会科学的概率为 0.375 $\left(=\frac{1\,500}{4\,000}\right)$，因为 4 000 名学生中有 1 500 学生的专业是社会科学。

假设你在体育馆碰到一名女生，并且想和她交谈。如果能准确地猜出她所学的专业也许会有助于实现这一目的。那么，她所学专业是自然科学的概率是多少？也就是说，假定知道她是一个女生，如何得出这个学生的专业是自然科学的概率？

要回答这个问题，首先注意提问“假定她是一个女生，那么这个学生的专业是自然科学的概率是多少？”与提问“从所有女生中随机选取学生的专业是自然科学的概率是多少？”是一样的。因而，我们先关注女生的总人数，表 11—6 中间的一行表示的就是这一内容。仔细观察这一行之后，我们得出的结论是 $\frac{600}{2\,100}=0.286$。换言之，2 100 名女生中有 600 人的专业是自然科学，而且因为假定每个女生出现在体育馆的概率相同，所以你碰到的学生所学专业为自然科学的概率是 28.6%。

你刚才在不知不觉地运用了贝叶斯法则。为了更正式地予以介绍，设 x 和 y 是两个随机事件。在之前的例子中，随机事件是指一个学生的性别和一个学生的专业。Prob (x') 表示的是 $x=x'$ 的概率（如一个学生是女生的概率），Prob (x', y') 表示 $x=x'$ 和 $y=y'$ 的概率（如一个学生是女生和学生专业是自然科学的概率），Prob $(y'|x')$ 表示在 $x=x'$ 的条件下 $y=y'$ 的概率（如根据学生是一个女生，她的专业为自然科学的概率）。贝叶斯法则对三种概率之间的关系作出如下表述：

$$\text{Prob}(y'|x')=\frac{\text{Prob}(x',y')}{\text{Prob}(x')} \tag{11.6}$$

这就是我们之前已经计算过的内容，即：

$$\text{Prob（学生是女生和她的专业是自然科学）}=\frac{600}{4\,000}$$

$$\text{Prob（学生是女生）}=\frac{2\,100}{4\,000}$$

这样，

$$\text{Prob（学生是女性，她的专业是自然科学）}=\frac{\frac{600}{4\,000}}{\frac{2\,100}{4\,000}}=\frac{600}{2\,100}$$

下面的特性可能有助于对贝叶斯法则的运用。假定 y 只有两个可能值：y' 和 y''。首先观察：

$$\text{Prob}(x')=\text{Prob}(x',y')+\text{Prob}(x',y'') \tag{11.7}$$

这一等式成立是因为相互排他性的存在。当 $x=x'$ 时，$y=y'$ 或者 $y=y''$，因为 y 只有这两个值。把公式（11.7）的右项代入公式（11.6）的分母中，得出：

$$\text{Prob}(y'|x')=\frac{\text{Prob}(x',y')}{\text{Prob}(x',y')+\text{Prob}(x',y'')} \tag{11.8}$$

转换一下公式（11.6）中 x 和 y 的作用，我们可以得到在 $y=y'$ 的条件下，分析 $x=x'$ 的概率，即：

$$\text{Prob}(x'|y')=\frac{\text{Prob}(x',y')}{\text{Prob}(y')}$$

左右两项都乘以 Prob (y')，即：

$$\text{Prob}(y') \times \text{Prob}(x'|y') = \text{Prob}(x', y') \tag{11.9}$$

对 $y=y''$ 也完成相同的步骤可以得出：

$$\text{Prob}(y'') \times \text{Prob}(x'|y'') = \text{Prob}(x', y'') \tag{11.10}$$

把公式（11.9）和公式（11.10）代入公式（11.8）的分母中，得出：

$$\text{Prob}\left(y'|x'\right) = \frac{\text{Prob}\left(y'\right)\times\text{Prob}\left(x'|y'\right)}{\text{Prob}\left(y'\right)\times\text{Prob}\left(x'|y'\right)+\text{Prob}\left(y''\right)\times\text{Prob}\left(x'|y''\right)}$$

这是一个用贝叶斯法则求出的计算 Prob ($y'|x'$) 的替代公式。

贝叶斯法则不仅可以用于你在体育馆与人见面做出推断的情形，也可以成为你参加电视节目《来做个交易吧》时的帮手。在这个节目中，每个选手需要从三扇门中选择一扇。其中有一扇门的后面藏有大奖（比如一辆崭新的汽车），另外两扇门的后面都是普通甚至是不受欢迎的奖品（比如一头坏脾气的驴）。虽然节目主持人知道每扇门后面的奖品是什么，但选手并不知道。在节目开始之前，汽车被随意放在一扇门的后面，所以选手起初认为大奖在 1 号门后面、2 号门后面和 3 号门后面的概率都是 $\frac{1}{3}$。

选手开始选择一扇门。因为她认为选哪扇门都一样，所以我们假设是 1 号门。但是，1 号在被打开之前，主持人先打开了另外两扇门中的一扇。现在，有一个关键的情况：主持人从不打开后面藏有汽车的门。也就是说，如果 2 号门后面有汽车，那么主持人会打开 3 号门；如果 3 号门后面有汽车，那么他会打开 2 号门；如果 1 号门后有汽车，那么可以打开 2 号或 3 号门（我们假定每扇门被打开的概率相同）。

假设主持人打开了 2 号门。然后他会提议：你想保留 1 号门还是选择 3 号门，而这一提议已经迷惑了许多的选手。选手做出选择之后，所有的门都被打开，这时选手知道自己得到了什么奖品。只有汽车更有可能藏在 3 号门的后面时，选手才会转而选择保留 3 号门。如果这个选手阅读过本章附录所讲的内容，那么她知道答案一目了然：转而选择 3 号。

为了证明这一推断，先分析汽车在哪扇门后这一无条件概率：

Prob（1 号门有车）$=\frac{1}{3}$，Prob（2 号门有车）$=\frac{1}{3}$，Prob（3 号门有车）$=\frac{1}{3}$，

要判定选手是否应该换门，有必要根据主持人已经打开 2 号门，计算出 1 号门有车的概率。这种情形下我们就可以使用贝叶斯法则，即：

$$\text{Prob}（\text{根据主持人打开 2 号门，1 号门有车}）= \frac{\text{Prob}（\text{主持人打开 2 号门,1 号门有车}）}{\text{Prob}（\text{主持人打开 2 号门}）} \tag{11.11}$$

分母可以表示为：

$$\begin{aligned}&\text{Prob}（\text{主持人打开 2 号门}）\\ &= \text{Prob}（\text{主持人打开 2 号门，1 号门有车}）+\text{Prob}（\text{主持人打开 2 号门，3 号门有车}）\end{aligned} \tag{11.12}$$

下面，将两项中的每一项放在一起计算：

$$\begin{aligned}&\text{Prob}（\text{主持人打开 2 号门，1 号门有车}）\\ &= \text{Prob}（\text{已知 1 号门有车，主持人打开 2 号门}）\times\text{Prob}（\text{1 号门有车}）\\ &= \frac{1}{2}\times\frac{1}{3}=\frac{1}{6}\end{aligned} \tag{11.13}$$

注意，在之前的计算中，如果 1 号门有车，那么主持人打开 2 号门的概率是 $\frac{1}{2}$，打开 3 号门的概率是 $\frac{1}{2}$。

接下来，注意如果 3 号门有车，那么根据主持人从不打开后面有大奖的门规则，他必须打开 2 号门。利用这一事实，我们得出：

$$\begin{aligned}&\text{Prob（主持人打开 2 号门，3 号门有车）}\\&=\text{Prob（已知 3 号门有车，主持人打开 2 号门）}\times\text{Prob（3 号门有车）}\\&=1\times\frac{1}{3}=\frac{1}{3}\end{aligned}\tag{11.14}$$

将公式（11.13）和公式（11.14）代入公式（11.12），得到：

$$\begin{aligned}&\text{Prob（主持人打开 2 号门）}\\&=\text{Prob（主持人打开 2 号门，1 号门有车）}+\text{Prob（主持人打开 2 号门，3 号门有车）}\\&=\frac{1}{6}+\frac{1}{3}=\frac{1}{2}\end{aligned}\tag{11.15}$$

最后，将公式（11.13）和公式（11.15）代入公式（11.11）：

$$\text{Prob（已知主持人打开 2 号门，1 号门有车）}=\frac{\frac{1}{6}}{\frac{1}{2}}=\frac{1}{3}$$

主持人打开 2 号门后，选手知道汽车要么在 1 号门后，要么在 2 号门后。因为我们已经证明 1 号门的无条件概率为 $\frac{1}{3}$，所以 3 号门的有条件概率为 $\frac{2}{3}$。因而选手应该总是要求换门。主持人回避 3 号门的事实也提供了一个信息。如果 3 号门有大奖，那么主持人必须回避这扇门，而这一举动更说明 3 号门有大奖。虽然以上的信息比较直观，但贝叶斯法则的魅力就在于告诉我们如何准确使用这一信息来更新我们的信念。



第
12
章

谎话与说谎者透露的信息：廉价谈话博弈

学习目标

12.1 博弈论中的交流
12.2 表明信息
12.3 表明意图

本章导读

谈话是廉价的，付出的只有气息。

——《随员》（*The Attaché*）

第 11 章讨论了在个人行为传递信息时只有自己知道自己意图的情形。通过赠送昂贵奢侈的礼物，追求者可以向一个女人表达自己的爱。通过长时间的工作，实习生可以证明自己有多勤奋。通过给顾客提供退款保证，公司可以向大众传递产品质量好的信息。在所有的这些例子中，行为传递了信息，因为选择行动的代价很高，并且随参与者身份不同而不同。一个钻戒代表真爱，因为只有真爱的人才会心甘情愿付出这么多。长时间的工作说明实习生有职业道德，因为本质上就很懒的人发现长时间的工作无趣且无意义。一个公司若知道他的产品是劣质的，就会发现退款代价太高，因为公司必须频繁地给顾客退款。

在这章中会提到许多信息传递的场景，为了使你更有兴趣，我们将做一个重要的调整：采取行动没有任何成本。说“没有任何成本”，意思是采取行动并不影响参与者（任何一个参与者）的收益。那么采取行动对所有的参与者来说，成本都是一样的，即都为 0。

尽管我们的第一反应可能是想弄清楚一个没有任何成本的行动是如何产生那么大的影响，但实际上，我们每天都采取零成本的行动来影响其他人的行为。这些零成本的行动神奇之处在哪里？就在于它是我们平时所讲的话。它可能是一个说服你买车的电视商业广告；可能是一个候选人竞选时的竞选承诺；可能是汽车修理工说的说服你的车需要修补的话；可能是一条发给朋友确定晚饭的时间和地点的短信，也可能是来自尼日利亚一个被废黜的将军妻子的邮件。

一个零成本的行动被称为消息。以消息为载体的博弈被称为廉价谈话博弈。一个重要的假设是发送错误的消息（“说谎”）没有成本；所以，我们不对消息的真假做假设。因为谈话是廉价的，一个人可能会想消息什么

时候才是有信息量的？解决这个问题是这章的首要目的。在分析师的建议下，你什么时候应该买股票？你是否应该对竞选人不提高税收的竞选承诺深信不疑？汽车修理工说你的车需要换一个变速器，你能信他的话吗？有信息量的消息对人们的行为有很大的影响，但如果它们缺乏实质性的内容，这些消息就是一堆废话。

消息在两个不同的场景中产生大的影响，即可以“传递信息”或“表明意图”。第一个场景就是在第 11 章中提到的：比如，一个参与者知道关于收益的信息，他发消息给其他参与者，这就可能影响其他参与者对收益信息的了解以及如何采取行动。第二个经典的场景和完全信息博弈有关。这里，一个参与者的消息并没有传递关于这场博弈的任何信息；所有人都知道博弈的特征。相反，这条消息传递的可能是这个参与者的预谋。这在多重纳什均衡博弈中尤其如此。比如，回忆第 4 章的电话博弈，在这场博弈中，科琳和温妮的电话谈话被切断，他们必须决定是等电话还是回电话。每个人都希望对方把电话打过来，但如果有事必须联系的话，每个人还是愿意主动回电话。在这场博弈中，有两种纯策略纳什均衡：科琳打回电话，温妮等电话；温妮打回电话，科琳等电话。第一次电话刚开始的交流可以帮助参与者选定其中的一种均衡。还存在一种对称性混合策略纳什均衡，在这种均衡中，每个参与者可以随机选择策略。尽管第二种均衡在平等对待参与者上有很大的优势，但它可能时不时地使合作失败。正如我们将要看到的，电话博弈之前的廉价谈话允许每个参与者表明自己的意图以便促成一个更好的合作结果。

在 12.1 小节中，我们思考在一个博弈论世界中（和现实世界有一点点不同），交流意味着什么。12.2 小节探索了当参与者试图传达信息时消息的内容。12.3 小节讨论了前博弈沟通在表明意图时所起的作用，我们也探索了理论预测如何与实验证据吻合。结果是人们在实际中的行为在很大程度上影响了允许参与者交流的预期效果。

12.1 博弈论中的交流

这个小节主要讲三点。第一，消息不仅仅靠语言传递；其他参与者能了解到的任何一个参与者所做的任何事都可以成为消息。第二，消息的信息内容不一定与消息的字面意思一致。一条消息的意义是被推出来的，而不是被假设的。第三，尽管一个参与者可能想要欺骗别人，但在纳什均衡点上就没必要说谎。欺骗的欲望并不一定导致欺骗行为的发生，但它可以阻止信息披露。在接下来的例子中会充分解释这几点。

什么是交流？交流涉及一个参与者给另一个参与者“发送消息”。一条消息有很多种形式。它可以是语言，可以是文字。如果你在纽约的一家中国餐馆想要点宫保鸡丁，你会用英语对服务员说“我要宫保鸡丁”。如果你在上海，也不懂汉语（服务员也不懂英语），你仍然可以用英语说“我要宫保鸡丁”，但这不可能达到预期的效果。然而，如果菜单是中英文的，你可以指着菜单上的宫保鸡丁菜名，“指”菜单这种身体语言也是传达消息的一种。

一条消息可以是一个参与者发送的且被另一个参与者理解的无任何成本的行动（就是说，通过一个人的感官意识到的），可以是口语（被听到）或涉及肢体语言（可以被看到或被感觉到）。在影片《骗中骗》(*The Sting*) 中，演员之间用手指头在鼻子上轻轻一点传递信息。类似的，轻微的手势被用在竞标者和拍卖者之间，表明竞标者已经打败其他竞标人接受拍卖结果。消息也可以是你挂在门上的悬挂物，这样舍友就知道你有“同伴”。消息的范围包括可以传递信息的一系列说的话和做的手势。

然而，消息并不总是能够传递信息。这就给我们带来了消息的意义这个命题。当然，语言有字面意思，查下字典就知道，但我们知道说话人有时候故意说一些晦涩、误导人甚至是欺骗人的话。一个人可以说和事实完全相反的话。要达到成功的交流，必须满足两个条件。第一个条件，参与者对信息的字面意义必须有一个共同的认知。比如，在美国，“摇头”这个手势的字面意思就是“不”。但在印度的一些地方，相同的肢体语言表示“是”的意思。我和来自印度的一个研究生（现在已成为一个超棒的博弈理论家）的第一次交流非常失败，因为我们缺乏那种共同的认知。

共同语言就是对消息字面意思的共同理解，是传递信息的必要但非充分条件。这个事实把我们引入第二个条件：那就是发送消息的人没有任何欺骗信息接收者的动机。在 1988 年竞选总统时，乔治·布什做出明确的承诺：“看着我的嘴唇，认真听我讲……没有任何新的税收！”尽管发音很清楚，这条消息传递出的对未来的承

诺却很模糊，因为选民质疑布什不管自己内心真实的信念是什么，他都可能做出这样的承诺。同期盖洛普民意调查显示有 68% 的选民对此持怀疑态度，他们认为布什如果当选一定会提高税收。很明显，选民读的是他的思想，而非嘴唇。

于是，当两个人有共同的语言且没有欺骗对方的动机时，信息就被成功传递。欺骗是这章要研究的核心概念。我们将对欺骗何时起作用与它如何影响消息所传递信息的内容感兴趣。对你来说，最重要的一点就是在这个阶段中，意识到消息的真正意义是博弈论需要解决的一个方面。

发音清晰的消息不一定有信息含量，那些不在正式语言范围内的消息却可能包含丰富的信息。为了突出说明信息为什么可以是任意的且有信息量的，不说话也可以传递信息。当美国国会就美国职业棒球大联盟使用类固醇的问题举行听证会，麦圭尔（McGwire）被以前的队友指控使用类固醇。他在这个问题上表示沉默就说明了一切。

既没有承认也没有否认，他使每个人都相信他服用了类固醇。他的不否认指控就说明了问题：如果他没有服用类固醇，那么最好的办法就是如实说“没有服用”。

这里，我们想要解决的最后一个问题是：说谎，说模棱两可的话，说搪塞的话的意义何在？根据韦氏词典，“说谎”就是“说一些不真实的话语达到欺骗别人的目的”。但有时候也不是这个字面上所说明的情况，有时候说谎者的意图可以被别人推断出来。让我举个我家的例子。我女儿在回答某些问题时，当真实答案是“不能”的时候，她的答案是“可能”。例如，如果我问她“你的家务活做完了吗？”她总是回答“可能做完了”。但我认为肯定是“没做完”，因为如果她做完了，她肯定会说“是的，爸爸，现在不要再烦我了。”当我的女儿说“可能做完了”的时候，她并没有说谎，因为我们都知道那是“没做完”的意思，也是我问题的真实答案。一个人必须成功的误导了他人才称得上说谎。

在博弈论模型中，我们将会探索到，人在很多情况下有欺骗的动机，他就会选择说谎。然而，在纳什均衡点上就没有人说谎，因为在纳什均衡点上，这个动机会被其他参与者考虑到。尽管在我们的分析中不允许任何一个参与者被骗，那并不意味着信息就是真实的：欺骗的动机阻止了信息的公开。如果一个侦探问一个嫌疑犯他是否认罪，他的答案不一定是真实的：如果他没有犯罪，他会说他不会认罪的；但如果嫌疑犯真的犯了罪，他也会说他没罪。这个侦探没有被欺骗，尽管他还对这件事情的真相一无所知。

在博弈论模型中说谎现象的缺失是令人遗憾的，因为在现实生活中，人们经常成功地欺骗了他人。遗憾的是，博弈论不能清楚地阐明说谎问题，因为我们的参与者都太聪明了。它不能解释关键的问题，比如在莎士比亚的《李尔王》(*King Lear*）中，李尔王是不列颠的一位年老的国王，他决定处置他的王位。他让他的三个女儿告诉他她们有多爱他。两个姐姐高纳里尔和丽甘奉承爸爸给了错误的答案，而小女儿科迪莉亚保持沉默，因为她无法用语言来表达对父亲的爱。李尔王就与科迪莉亚断绝父女关系而把王位传给了高纳里尔和丽甘。不久他就明白了那些奉承的话都是谎言，小女儿的沉默才是真爱。现在，在博弈论纳什均衡点上，李尔王可能不会了解每个女儿的爱到底有多少，但他也不会被骗。

12.2 表明信息

一个廉价谈话博弈有三步：

步骤 1：“自然”选择信息发送者的类型。
步骤 2：信息发送者了解她的类型并选择消息。
步骤 3：信息接收者查看发送者的消息，在这条信息基础上修正他对发送者的观点并采取某种行动。

这条消息成为廉价谈话博弈而不是其他任何类型的信号传递博弈是因为由发送者选择的消息既不影响发送者的收益，也不影响接收者的收益。他们的收益只取决于发送者的类型和接收者的行动。当然，一条消息可以间接地影响收益，通过改变接收者对发送者类型的判断来影响接收者的行动。为了使这种情况发生，尽管发送者的类型一定会影响接收者的收益而不只是发送者的收益，在我们的例子中，我们将假设这种情况。

○案例研究

防御性医疗

廉价谈话的共同特征是何时信任专家的建议：一家

汽车维修店说你的车需要一次大的维修；你的股票经纪人建议你买一只股票；你的私人医生建议你做一次诊断测试。你什么时候应该听从这些建议？专家什么时候表达的是他所了解的真实情况？什么时候他是站在自己立场上而不是你的立场上劝你去做某些事？

思考一位有特殊症状的患者去咨询他的私人医生。医生检查完患者后要决定是否建议患者做一个很昂贵的医疗测试，比如磁性共振成像（MRI）。如果他不建议，患者就必须自己做出某种决定。患者意识到这个测试很有必要，但同时又觉得对他个人来说太贵，因为做这个测试的费用保险不能报销或者还需要占用他工作的时间，又或者做这个测试会使他不舒服。医生关心患者，但也许并非全是出于关心。她可能推荐患者做这个测试，即使她知道患者做了这个测试也不一定会好转，因为这样做只是避免了潜在的医疗事故诉讼。患者也意识到他的医生将会采取防御性医疗行为的可能性。

图 12—1 所示是贝叶斯博弈扩展形式。“自然”先行决定患者做这个测试的价值。有$\frac{1}{3}$的概率做这个测试对患者的病情有用；有$\frac{2}{3}$的概率做这个测试就是无用的。测试价值为多少只有医生在检查了患者之后才能知道。在知道对患者来说测试价值为多少之后，医生才决定是否建议患者做这个测试。如果她不建议做，那么由患者自己决定要不要做这个测试。如果她建议做，那么由患者决定是否遵从医生的建议。

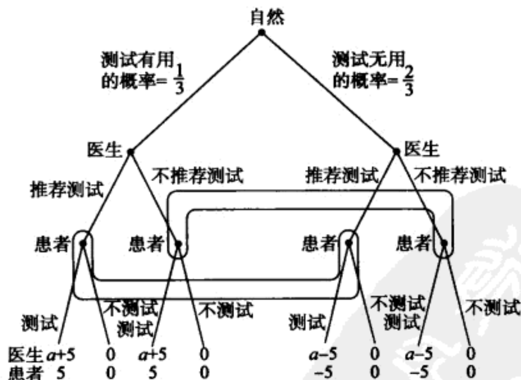


图 12—1　医生与患者的博弈

如果患者接受测试，当测试对患者的病情有用时，患者得到的收益为 5，反之为−5。如果他不接受测试，那么他的收益为 0。只有当测试对病情有用时，患者才做这个测试，虽然想弄清这一点，需要核实医生建议的真实性。医生从做测试中得到的收益为$a+v$，这里v是患者从测试中得到的收益，且$a \geqslant 0$。如果测试对患者病情有用，$v=5$；如果没用，那么$v=-5$。a是从医疗事故方面来说，医生从测试中得到的收益。如果$a=0$，那么医生和患者的收益相同；所以，当且仅当测试对患者有用时，医生才会建议做这个测试。然而，如果$a>0$，那么就会有利益争斗，因为即使当测试对患者无用时，医生也愿意建议患者做这个测试。最后，假设如果没有做这个测试，医生的收益为 0。

首先注意存在一个混同精炼贝叶斯纳什均衡，在这个均衡中，患者不应该相信医生：

- **医生的策略：**不管是不是对患者有用，都建议做这个测试。
- **患者的策略：**忽略医生的建议。
- **患者的信念：**不论医生是否建议做这个测试，这个测试对他的病情有用的概率只有$\frac{1}{3}$。

这个均衡就是所谓的胡说均衡：这个医生的信息并不比无聊的唠叨更有用。在廉价谈话博弈中，一个胡说均衡就是一个混同均衡。假设这个医生的策略使她做出相同的建议——“做这个测试”，不管她的类型是什么，她的信息都是空洞无用的。那么患者会坚持他一贯的信念，因为这并没有对他先前的信念产生任何影响。在所有这些信念的影响下，当患者被建议做测试时，他做测试的收益就为：

$$\left(\frac{1}{3}\right)\times 5+\left(\frac{2}{3}\right)\times(-5)=-\frac{5}{3}$$

这个收益低于他不接受做测试的收益。因为患者本来不打算做这个测试，不管医生是否建议做这个测试（其实两种结果是一样的），医生的收益都为 0。这个均衡没有产生令人激动的效果是因为医生的建议是没有任何价值的（因为建议并不是建立在患者的真实情况上）。当然，这个建议就会被患者忽略。

实际上，每场廉价谈话博弈都存在一个胡说均衡。如果接受信息的人认为发送者的消息和发送者的类型无关，那么他会忽略这条信息，只用自己的先验信念决定如何采取行动。但如果接收者忽略发送者的消息，那么发送者就会不在乎说什么，因为这并不影响接收者的行

为（所有的消息成本都为 0）。



点拨

对于每一场廉价谈话博弈，总存在一个胡说均衡。

在这个有趣的命题中，是否也存在这样一个均衡，在这个均衡中，消息实际上是包含信息量的。让我们接下来思考满足分离精炼贝叶斯纳什均衡的条件：

- **医生的策略**：当且仅当测试对患者有用时，建议患者做这个测试。
- **患者的策略**：遵从医生的建议。
- **患者的信念**：如果医生建议做测试，那么测试有用的概率为 1。
  如果医生不建议做测试，那么测试无用的概率为 1。

这个策略组合会导致一个令人满意的结果，因为医生总是为患者考虑，而患者也相信医生。我们需要评估的是它什么时候是一个均衡状态。患者的信念是连贯的且应该是显性的。（如果是隐性的，你应该和你的指导员安排一下时间。）假设只有当测试对患者是有用的，医生才建议患者做测试，患者的策略明显就是最优的。这个结果就使我们评估医生策略的序贯理性。假如，实际上这个测试是有用的。那么医生从建议患者做测试中得到的收益为 $a+5$，因为，根据患者的策略，患者会遵从医生的建议。既然医生从不建议患者做测试中得到的收益为 0，当且仅当 $a+5 \geq 0$ 时（因为 $a \geq 0$，所以 $a+5 \geq 0$），医生的策略为最优策略。如果这个测试对患者有用且无医疗事故的发生，医生肯定会建议患者做测试。具有争议的是当测试对患者病情无用的情况。那么这个策略就使医生做出不建议的决定，因为她的收益为 0。而建议做测试得到的收益为 $a-5$，所以当 $0 \geq a-5$，即 $a \leq 5$ 时，医生不建议做测试是最优策略。

当 $a \leq 5$ 时，存在一个有真实建议的均衡。相反，当 $a>5$ 时，只存在一个胡说均衡，所以医生的建议就是无任何信息含量的。这里存在一个普遍的原则。当 $a=0$ 时，医生和患者的利益相吻合，因为只有当测试使患者情况变得更好时，医生才建议患者做测试。然而，当 a 逐步增大时，医生和患者的关注点就会有分歧，因为医生的建议越来越受到避免医疗事故的影响，而患者越来越关心测试的有效性。所以，当 a 比较小时（即 $a \leq 5$），医生主要关心患者的健康和她的建议是否有效。相反，当 a 比较大时（即 $a>5$），医生主要关心医疗事故的问题。

一般来说，两方的利益越一致，越存在一个均衡，在这个均衡中信息是有效的。当他们的利益完全一致时，对发送者有利的就对接收者有利，在这种情况下，发送者没有误导接收者的动机；实际上，它可能会误导自己。当利益不是很一致时，只要没有太大的分歧，这种直觉就一直奏效。现在思考利益完全相反的情况，如在零和博弈中的情况。在那种情况下，对发送者越有利的对接收者就越不利。发送者就试图误导信息接收者，在这种情况下，信息接收者就不应该相信发送者的消息。既然在均衡中，没有人被误导，结果就是不应该相信信息。更直观地说，如果利益完全相反，那么消息就是没有任何信息量的。



点拨

信息发送者和接收者的利益越一致，消息就越有可能是有信息量的。当利益足够接近时，存在一个均衡，在均衡中，消息是有信息量的。当他们的利益很不一致时，所有的均衡中消息都是没有信息量的。

12.1 理解测验

在如图 12—2 所示的廉价谈话博弈中，求出两个分离均衡和两个胡说均衡。

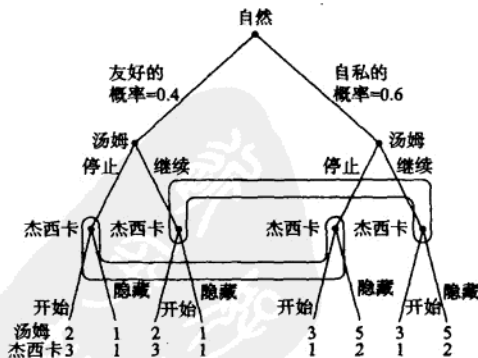


图 12—2

购买股票的建议

近年来，证券业一个有争议的话题就是提供投资建议和从事投资银行业务引起的利益冲突。在任何时候，像星巴克这样的公司会被诸如美林这样的投资证券公司吞并，美林公司的分析师将为它们的客户提供“买进”、“持有”、“卖出”的建议。如果星巴克想要通过发行股票来聚集资金，他可能会雇用投资银行公司。因为星巴克当然不想听到抛售自己股票的建议，为了揽到星巴克投资银行的业务，美林公司会使他的分析师不做出卖出星巴克股票的建议吗？

为了探索这个问题，思考图 12—3 中描述的博弈。“自然”先决定投资分析师根据自己的研究结果，是否相信这支股票是胜于、劣于还是和股市上的其他股票一样的表现。假设每个事件发生的概率为$\frac{1}{3}$。如果一个投资分析师打算投资自己的钱，当它胜于其他股票时，他将买入；当它和其他股票持平时，他继续持有；当股票预期表现不佳时，他就卖出股票。一旦了解了股票的预期表现，分析师就发布投资建议：买进、持有或者卖出。这个建议就是廉价谈话。投资者或客户（这场博弈中的信息接收者）了解了分析师的建议，尽管他并不知道投资分析师的真实信念。她将由分析师的建议推断并更新自己关于股票表现的信念，然后决定是买进、持有还是卖出（假如她已经拥有一些股票）。

表 12—1 列出了分析师和投资人的收益与股票的真实表现和投资人做法之间的关系。投资人从采取最好的做法中得到 1 个收益，最好的做法是当股票预期表现良好时买进股票；当股票预期一般时，继续持有股票；当预期表现不佳时，卖出股票。从采取最坏的做法中得到的收益为−1，最坏的做法是当股票预期表现良好时卖出股票和当股票预期表现不佳时买进股票。还有就是收益为 0 的情况。当投资者买进时，分析师的收益等于投资者的收益加 a；当投资者卖出时，分析师的收益等于投资者的收益减去 b，这里 $a>0,b>0$。分析师的收益值随投资者的收益值变化而变化反映了当他客户的投资组合表现更好时，分析师得到的报酬就更高。

表12—1　买卖股票建议博弈中的收益

状态	行动	分析师的收益	投资者的收益
表现良好	买进	$a+1$	1
表现良好	持有	0	0
表现良好	卖出	$-b-1$	−1
表现一般	买进	a	0
表现一般	持有	1	1
表现一般	卖出	$-b$	0
表现不佳	买进	$a-1$	−1
表现不佳	持有	0	0
表现不佳	卖出	$1-b$	1

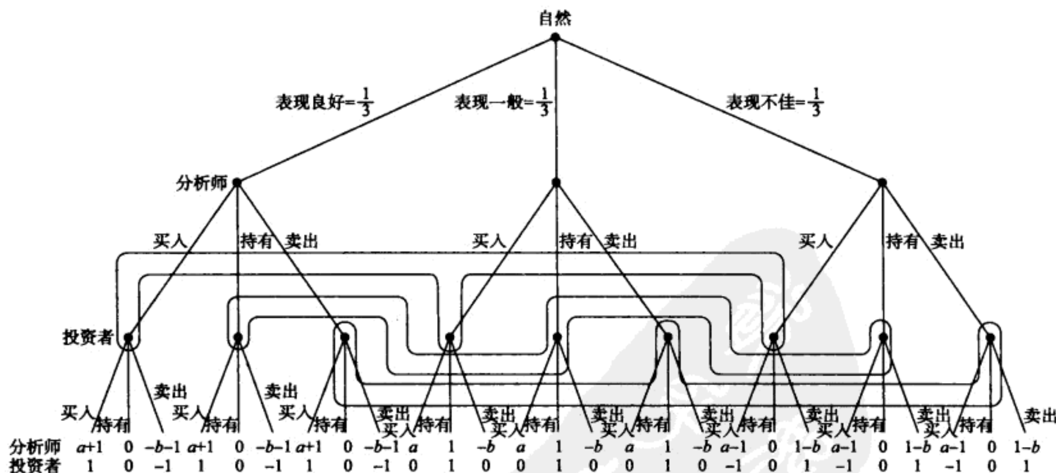


图 12—3　分析师与投资者之间的博弈

这个附加项是加 a 还是减 b 取决于投资银行的方案。当客户被引导卖出股票时，分析师（和他的公司）的利益就受到损害；当客户被引导买进股票时，分析师（和他的公司）就从中受益。和廉价谈话博弈一样，收益值并不仅仅取决于分析师的建议。

通过思考策略组合和信念使建议完全有效，让我们开始我们的分析。在这个场景中，分析师给了中肯的建议且投资者相信建议并遵从这个建议采取行动：

- **分析师策略：**
  当股票表现良好时，买入股票。
  当股票表现一般时，继续持有股票。
  当股票表现不佳时，卖出股票。
- **投资者的策略：**遵从分析师的建议。
- **投资者的信念：**
  当分析师建议买进时，股票上涨的概率为 1。
  当分析师建议继续持有时，股票表现一般的概率为 1。
  当分析师建议卖出股票时，股票下跌的概率为 1。

投资者的信念和分析师的策略利益的一致性应该很明显了。给定投资者的信念，也不难证明投资者的策略是最优的。唯一不确定的是分析师的策略是不是最优的。假设分析师认为股票将会上涨，那么他的策略就会使他提出买进的建议，产生 $a+1$ 个收益，因为它引导投资者采取正确的行动（这里分析师附加了 1 个收益），买入股票（分析师预计收益为 a）。相反，如果他做出继续持有的建议，那么分析师的收益为 0，如果做出卖出的建议，收益会更低，为 $-b-1$。那么，很明显，当分析师认为股票将上涨时，提出买进建议是最优策略。

最后，思考股票表现不佳时的情况。在这种情况下，分析师的策略使他做出卖出的建议，产生收益 $1-b$。因为给出正确的建议使他得到 1 个收益，使投资者卖出股票又使他损失 b 个收益。作为对比，做出持有股票的建议使他得到的收益为 0，建议买进时得到收益为 $a-1$。达到均衡必须满足 $1-b \geq 0$（所以卖出比继续持有要好）和 $1-b \geq a-1$（所以卖出比买进好）。这两个不等式就等同于 $1 \geq b$ 和 $2 \geq a+b$。

总之，我们发现对所有的股票类型来说，要使分析师的策略为最优策略必须满足三个条件：$a \leq 1, b \leq 1$ 和 $a+b \leq 2$。因为当前两个条件满足时，第三个条件肯定满足，所以当 $a \leq 1$ 和 $b \leq 1$ 时，这是一个精炼贝叶斯纳什均衡。如果投资银行的组合很重要，即如果引导股票买入的利益足够多或者引导股票卖出的损失足够大（或两者都有），那么 a 或 b（或者两者）的值大，所以对建议完全有效来说，它不是一个纳什均衡。然而，如果分析师在很大程度上在乎客户的利益，a 和 b 的值就小，那么投资者就可以相信分析师提出的建议。图 12—4 描述了 a 和 b 满足什么样的条件使建议完全有效是一个纳什均衡。

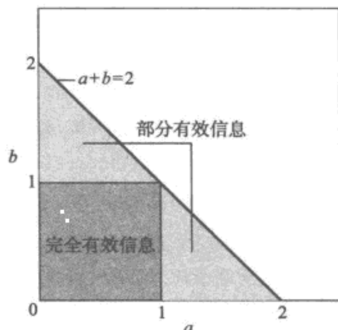


图 12—4　股票建议的信息内容

现在假设 $b>1$，那么引导客户卖出股票的损失是非常大的。分析师总是在透露事实时，就永远不是一个纳什均衡。但分析师的建议可能至少是部分有效的吗？为了解决这个问题，思考下面的半分离策略组合：

- **分析师的策略：**
  当股票表现良好或表现一般时，建议买进。
  当股票表现不佳时，建议继续持有。
- **投资者的策略：**
  当分析师建议买进就买进。
  当分析师建议继续持有或卖出时，卖出股票。
- **投资者的信念：**
  当分析师建议买进时，认为股票上涨的概率为 $\frac{1}{2}$，股票表现一般的概率为 $\frac{1}{2}$。
  当分析师建议继续持有时，认为股票表现不佳的概率为 1。
  当分析师建议卖出时，认为股票表现不佳的概率为 1。

从投资者的信念开始分析。因为当股票表现良好或表现一般时，分析师做出买进的建议，所以投资者应当认为股票表现不佳的概率为 0。假设在每种情况下，表现良好和表现一般的预期概率都为$\frac{1}{3}$，根据贝叶斯定律，每个的后验概率为：

$$\frac{1}{2}=\frac{\frac{1}{3}}{(\frac{1}{3})+(\frac{1}{3})}$$

相反，当分析师做出继续持有的建议时，投资者认为股票表现不佳的概率为 1 是正确的，因为只有当股票预期表现不佳时，分析师才做出继续持有股票的建议。那么投资者的信念就和分析师做出的买进或继续持有的建议一致。当建议卖出时，这不会对投资者的信念产生影响，因为根据分析师的策略，他不会做出卖出的建议。在这种情况下，我们已经假设投资者相信股票将表现不佳。

假设投资者的信念具有连贯性，接下来我们分析她的策略。据说当建议买进时，他就应该买进。这样做她的预期收益为：

$$\frac{1}{2}\times 1+\frac{1}{2}\times 0=\frac{1}{2}$$

如果股票表现良好的收益为 1，如果表现一般的收益为 0。继续持有股票的预期收益为：

$$\frac{1}{2}\times 0+\frac{1}{2}\times 1=\frac{1}{2}$$

卖出股票的预期收益为：

$$\frac{1}{2}\times(-1)+\frac{1}{2}\times 0=-\frac{1}{2}$$

那么买进股票确实是最优策略（尽管继续持有的策略使她得到相同的收益）。当分析师建议继续持有或卖出时，投资者相信股票将下跌，她的策略使她做出卖出的决定。那么对分析师的所有建议来说，这就是投资者的最优策略。

最后，思考分析师的策略。如果股票将上涨，他做出买进的建议获得的收益为 a+1，这好过做出继续持有或卖出建议获得的收益 $-b-1$。当股票表现平平，他的策略也使他做出买进的建议，尽管这次的收益只有 a。相反，如果他建议继续持有，得到的收益更低，为 $-b$。最后，如果股票将下跌，做出继续持有（或卖出）的建议产生的收益为 $1-b$，因为它引导投资者卖出股票。这个收益必须比 $a-1$（做出买入建议时得到的收益）大。所以，我们需要使 $1-b \geqslant a-1$ 或 $a+b \leqslant 2$。

只要 $a+b \leqslant 2$，当股票上涨或表现一般时分析师做出买进的建议或当股票将下跌做出卖出的建议，且当股票上涨时投资者买进，当被建议继续持有或卖出时，投资者做出卖出的决定就是一个均衡状态。注意条件式 $a+b \leqslant 2$ 比完全分离均衡的条件式（$a \leqslant 1$ 和 $b \leqslant 1$）弱，满足了 $a \leqslant 1$ 和 $b \leqslant 1$，就一定得出 $a+b \leqslant 2$。当然，如图 12—4 所示，如果 $a>1$ 或 $b>1$，也有可能 $a+b \leqslant 2$。比如，如果 $a=0.6$ 和 $b=0.8$，那么分离均衡存在，因为 $a \leqslant 1$ 和 $b \leqslant 1$。但是如果从引导投资者卖出股票的建议使投资银行失去更多的业务，比如现在假设 $b=1.2$，那么分离均衡就不存在（因为 $b>1$），但存在半分离均衡（因为 $a+b=0.6+1.2 \leqslant 2$）。

当 b（或 a）的值越高，分析师和投资者的利益分歧就越大，因为分析师更关注的是一个公司股票的买卖向他所在公司投资银行业务的影响。这使得分析师很难向他的客户提供真实的建议。当然，他不会欺骗，比如投资者知道继续持有股票的建议就意味着她应该卖出股票，但建议的信息量就会下降。现在，一个买进的建议表明股票或者上涨或者表现一般，且投资者不知道哪一个是真的。

分析师建议的信息内容在华尔街已经具有法律意义。表 12—2 报告了对这一课题早期研究的发现。建议的不对称性令人惊奇：尽管超过 15% 的建议是强买，实际中没有一个强卖的例子，并且几乎 95% 的建议是继续持有、买入或者强买。

表12—2　实际的股票建议

建议	次数	累计百分比
强买	38	15.2%
买	128	66.4%
继续持有	70	94.4%
卖	14	100.0%
强卖	0	100.0%

更有意思的是股票建议的意义发展到哪一个程度的时候，投资者开始了解到一个继续持有的建议的真正含义是卖出的建议。

随着语言继续发展，引入的新的术语如“掉期交易”和“撤销”。这两个是证券公司最不愿意做出的建议，证券公司没有明说，它们实际上就意味着“卖出”。正如瓦尔特·司各特（Walter Scott）两个世纪前说的那样，“当我们第一次欺骗别人时，我们会说：‘我们织了多么复杂的网！’”

12.2 理解测验

图 12—5 所示的廉价谈话博弈中，考虑一个半混同策略：如果莱斯莉为类型低或类型中时，她立刻发消息，如果为类型高时，她稍后发消息；如果立刻发消息，加里选择行动 a，如果稍后发消息，加里选择行动 b。证明在指定的信念下，这是一个精炼贝叶斯纳什均衡。

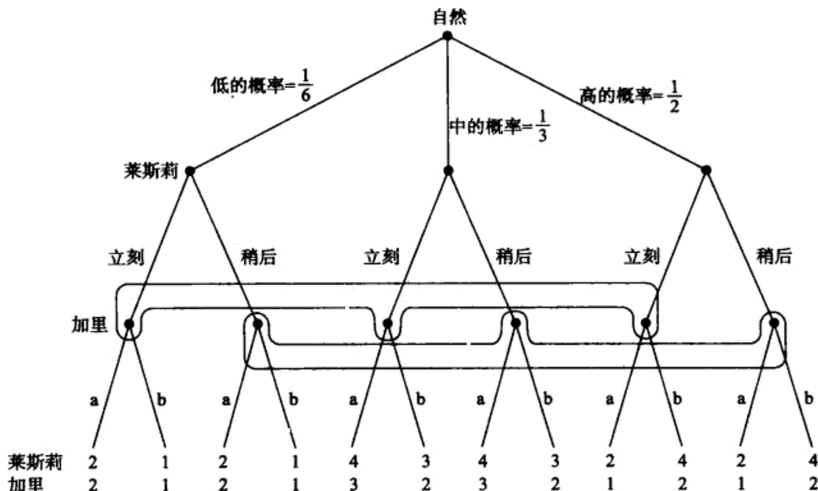


图 12—5

12.3 表明意图

前面的小节已经讨论了不完全信息博弈，在不完全信息博弈中，消息传递所有参与者面临环境的信息。现在我们假设环境是所有参与者的共同知识，即使所有参与者被期望做出均衡策略的选择，每个参与者仍然不知道如何做。为了探索信息在加强成员间合作中所起的作用，在这个战略场景中再增加一个阶段——**前博弈沟通**（preplay communication），在这个环节中，博弈的双方都有机会通过廉价谈话消息来传递他们的意图。

12.3.1 从理论上分析前博弈沟通

前面的例子证明了在开始采取行动之前，一个参与者的私人信息是如何传递给另一个参与者的。然而，这些并不是前博弈沟通可能有用的唯一场景。思考图 12—6 中的完全信息博弈。马特和菲奥纳正决定是看一部动作电影如《角斗士》(*Gladiator*)，还是去看肥皂剧如《十日恋爱日记》(*How to lose a guy in ten days*)。尽管马特酷爱动作电影而菲奥纳沉迷肥皂剧，但每个人最在乎的是一起去看一部电影。马特最期望的结果是他们都去看动作片，而菲奥纳最期望的结果是两人一起看肥皂剧。假设这个计划是他们下班后在电影院见面。然而，因为他们事前没有达成一个协议，每个人都要决定去看哪部电影，同时希望对方也做出相同的选择。

		菲奥纳	
		肥皂剧	动作片
马特	肥皂剧	2,3	0,−1
	动作片	1,0	3,2

图 12—6 看电影博弈

在这场博弈中存在两个纯策略纳什均衡。一个是他们两个都去看肥皂剧，在这种情况下，菲奥纳得到她的最高收益 3，马特得到的收益为 2；另一个是他们都去看动作片，这样马特实现他的最高收益。尽管这两种结果都是均衡状态，但马特和菲奥纳是否会在去看一部电影上达成一致还是个问题。基于这个理由，一些博弈理

论家发现混合策略均衡是一个更好的解决办法。它不仅不偏爱参与者中任何一方，也解释了现实生活中两个人不配合去看同一部电影的可能。

使用第 7 章提到的方法，让我们推导菲奥纳采取混合策略使得马特觉得选择两种纯策略中的任意一种，结果对他来说都一样的条件。令 f（对菲奥纳来说）表示菲奥纳去看肥皂剧的概率，那么当：

$$f\times 2+(1-f)\times 0=f\times 1+(1-f)\times 3$$

时，马特从选择看肥皂剧中得到的期望收益（表达式的左边）和他选择看动作片得到的期望收益（表达式的右边）是一样的。解关于 f 的方程式，我们得到 $f=\frac{3}{4}$。所以，如果菲奥纳去看肥皂剧的概率为 75% 时，两种选择对马特来说结果都一样。

当然，这个随机结果使菲奥纳满意了。如果 m（对马特来说）表示马特去看肥皂剧的概率，那么当：

$$m\times 3+(1-m)\times 0=m\times(-1)+(1-m)\times 2$$

时，菲奥纳从选择看肥皂剧中得到的期望收益和她选择看动作片得到的期望收益是一样的。解关于 m 的方程式，我们得到 $m=\frac{1}{3}$。

存在一个混合策略纳什均衡，其中马特去看动作片的概率为 $\frac{2}{3}$，菲奥纳去看肥皂剧的概率为 $\frac{3}{4}$，那么他们最终去同电影院看同一部电影的概率为：

$$\frac{3}{4}\times\frac{1}{3}+\frac{1}{4}\times\frac{2}{3}=\frac{5}{12}=0.42$$

这里，$(\frac{3}{4})\times(\frac{1}{3})$ 表示他们都去看肥皂剧的概率，$(\frac{1}{4})\times(\frac{2}{3})$ 表示他们都去看动作片的概率。这个均衡点上的预期收益为如下：

马特：

$$\frac{3}{4}\times\frac{1}{3}\times 2+\frac{3}{4}\times\frac{2}{3}\times 1+\frac{1}{4}\times\frac{1}{3}\times 0+\frac{1}{4}\times\frac{2}{3}\times 3=1.5$$

菲奥纳：

$$\frac{3}{4}\times\frac{1}{3}\times 3+\frac{3}{4}\times\frac{2}{3}\times 0+\frac{1}{4}\times\frac{1}{3}\times(-1)+\frac{1}{4}\times\frac{2}{3}\times 2=1$$

因为两人多数情形下不在同一剧院，所以他们得到期望收益都低于同看一部电影的最小收益 Z。

这里有个复杂的代数问题需要我们去求解：证明一点点的交流是如何有利于合作的。首先，注意如果菲奥纳和马特使用混合策略均衡，他们大于一半的概率会选择去不同的电影院（准确来说，是 58%）。现在假设我们允许他们在决定去哪个电影院之前通过发短信来交流。

我们头脑中的博弈如图 12—7 所示。菲奥纳和马特同时决定发短信“我们去看肥皂剧吧”，或者“我们去看动作片吧”。在各自收到短信之后，每个人决定到底去哪里。

这里没有私人信息，所以子博弈精炼纳什均衡是一个更合适的解决方案。存在许多这样的均衡，包括短信是无效的均衡，但我们还是关注信息有效的均衡。思考下面的策略组合：

- **马特的策略：**

  发送短信“让我们去看肥皂剧吧”的概率为 M，发送短信“让我们去看动作片吧”的概率为 $1-M$

  如果他们都发送短信“让我们去看肥皂剧吧”，那么去看肥皂剧。

  如果他们都发送短信“让我们去看动作片吧”，那么去看动作片。

  如果发送不同的短信，那么去看肥皂剧的概率为 $\frac{1}{3}$。

- **菲奥纳的策略：**

  发送短信“让我们去看肥皂剧吧”的概率为 F，发送短信“让我们去看动作片吧”的概率为 $1-F$

  如果他们都发送短信“让我们去看肥皂剧吧”，那么去看肥皂剧。

  如果他们都发送短信“让我们去看动作片吧”，那么去看动作片。

  如果发送不同的短信，那么去看肥皂剧的概率为 $\frac{3}{4}$。

当他们都发送同样的短信时，这个策略组合使他们妥协看同一部电影。在他们发送不同的短信的情况下，他们仅仅采取我们之前推导的混合策略纳什均衡。通过后面归纳的方法，我们需要思考 4 个最终的子博弈。首先，思考两者都发送短信“让我们去看肥皂剧吧”的情况。那么他们的策略使他们都去看肥皂剧，实际上这是一个纳什均衡。在子博弈中，如果他们都发短信“让我们去看动作片吧”，那么他们都去看动作片，而且这也是一

个纳什均衡。在两种子博弈中，他们发送的短信不一样，马特去看肥皂剧的概率为$\frac{1}{3}$，菲奥纳去看肥皂剧的概率为$\frac{3}{4}$，我们已经知道这个子策略组合是一个纳什均衡。

将每一个子博弈替换为在均衡点上的收益，我们推导出图 12—8 中的博弈。而且，它使我们更容易推导图 12—9 中的战略形式。在这场博弈中，我们需要求解 M 和 F 的均衡值。

一般情况下，M 必须使得菲奥纳觉得她的两种可行的短信发送什么都无所谓，即：

$M\times3+(1-M)\times1=M\times1+(1-M)\times2$

解这个方程式，得到 $M=\frac{1}{3}$。F 的值必须使得马特觉得发什么短信无所谓，即：

$F\times2+(1-F)\times1.5=F\times1.5+(1-F)\times3$

解这个方程得到 $F=\frac{3}{4}$。所以，当 $F=\frac{3}{4}$和 $M=\frac{1}{3}$时，上述策略组合是一个子博弈精炼纳什均衡。

在这个均衡点上，让我们计算菲奥纳和马特最终去同一个电影院的概率。有两种情况，一种是他们发送同样的短信（在发送短信上他们合作），或者他们发送不同的短信，但他们还是去了同一个电影院。前一个事件发生的概率为（$\frac{3}{4}$）×（$\frac{1}{3}$）+（$\frac{1}{4}$）×（$\frac{2}{3}$），即$\frac{5}{12}$，则他们发送的短信不一样，发生的概率为 1-（$\frac{5}{12}$），即$\frac{7}{12}$，当他们商量好选择去哪里时，他们最终去同一个电影院的概率为$\frac{5}{12}$（之前推导的）。那么他们发送的短信不一样，但还是去了同一个电影院的概率就为（$\frac{7}{12}$）×（$\frac{5}{12}$），即$\frac{35}{144}$。他们最终去同一个电影院的概率就为这两种情况之和，即$\frac{5}{12}+\frac{35}{144}=\frac{95}{144}\approx0.66$。

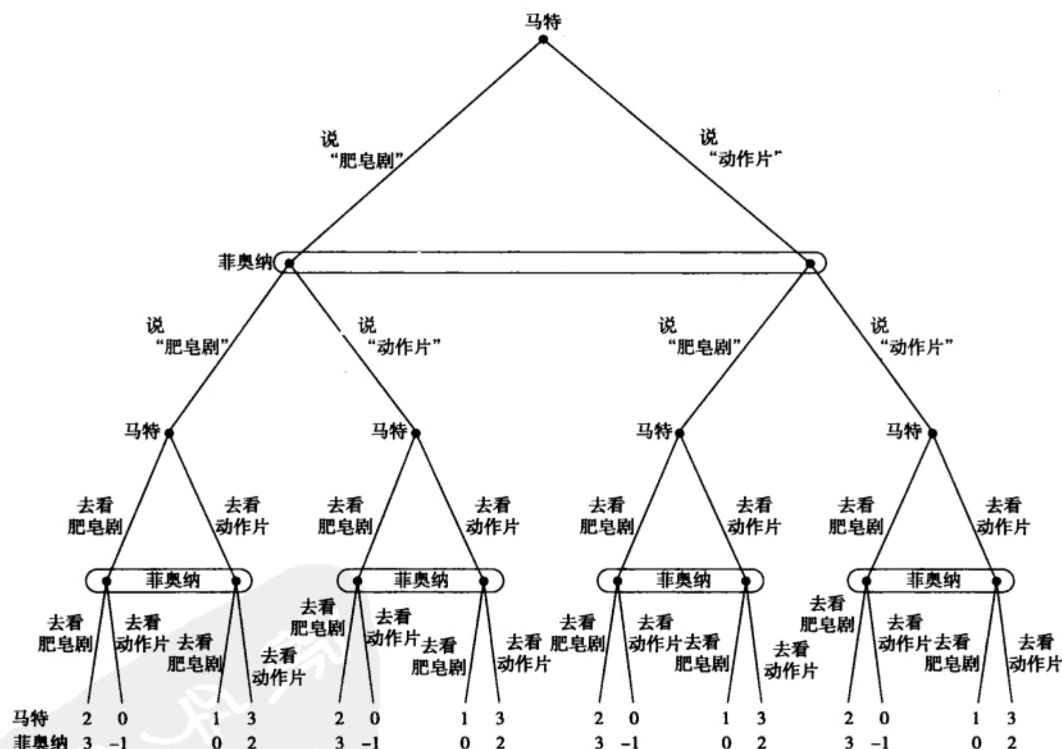


图 12—7　看电影前博弈沟通

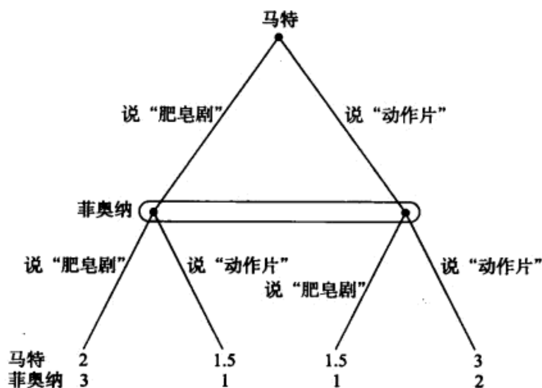


图 12—8　解决了最后的子博弈后的看电影博弈

		菲奥纳	
		“让我们去看肥皂剧吧”	“让我们去看动作片吧”
马特	“让我们去看肥皂剧吧”	2,3	1.5,1
	“让我们去看动作片吧”	1.5,1	3,2

图 12—9　图 12—8 博弈中范式表达的战略形式

当没有机会进行前博弈沟通时，菲奥纳和马特最终去同一个电影院的概率 0.42。一开始允许他们发送信息把概率提高了 0.66。这样，即使在完全信息博弈中，前博弈沟通还是很有帮助的。它使参与者有机会表达他们的意图，这样能帮助参与者协商他们的行动。这个场景中参与者确实有共同的兴趣是很重要的，这样的话，他们就有想要合作的基础。然而，如果菲奥纳想要和马特分手，那么菲奥纳的短信就是无效的信息。

12.3.2　实际情境中前博弈沟通

尽管前面的分析表明一轮交流可能会对合作有帮助，观察实际行为却更有说服力。接下来让我们回忆一些实验证据。由这个试验的收益（如图 12—10 所示）可知，当没有进行前博弈沟通时，混合策略纳什均衡使得两个人合作的概率为 37.5%（每个参与者选择 x 的概率为 0.25）。正如我们刚刚模拟的含前博弈沟通的博弈，但再附加第三条消息：每个参与者可以表明他采取 x 的意图，或者采取 y 的意图，或者什么也不说。

		参与者2	
		x	y
参与者1	x	0,0	2,6
	y	6,2	0,0

图 12—10　男女之间的博弈

参与者期望从均衡中得到的预期为：（1）当两者都表明采取 x 的意图，且实际都采取 x 或者其中一个参与者表明采取 x 的意图，而另一个表示沉默；（2）当两者都表明采取 y 的意图，且实际都采取 y 或者其中一个参与者表明采取 y 的意图，而另一个表示沉默；和（3）对任何信息，两者都采取混合策略均衡（具体地说，一个说采取 x，另一个说采取 y，或者两者都沉默）。在这个均衡中，参与者合作的概率为 49.9%。

大学生在实验室的行为就非常接近这个理论的预测。如果没有前博弈沟通，预测可能合作的概率为 37.5%；而实际上合作的概率为 41%。预期前博弈沟通会使合作的频率增加 33%=（0.499−0.375）/0.375，从 37.5% 增加到 49.9%。实际上，在试验中增加了 34%=（0.55−0.41）/0.41，从 41% 增加到 55%。我们所有的预测都是如此准确！

尽管前博弈沟通是很有用的，非合作的概率也有 45%。能保证产生合作的结果是有效的是在前博弈沟通阶段只允许一个参与者发送信息。在试验中，被允许交流的参与者几乎总是在她的最优策略下表明她的意图；比如，如果菲奥纳有权利发送短信，那么她会发送去看肥皂剧的信息来表明她的意图。单向交流被证明是最高效的，它可以使得合作的概率达到 95%！

图 12—11 猎鹿博弈中也探讨了前博弈沟通的作用。这里有两种纯策略纳什均衡：（a,a）和（b,b），两个参与者都认为后者比前者好。（b,b）的确是这场博弈的最好结果。对比这场博弈和男女之间的博弈，在这两场博弈中，参与者支持的均衡策略完全不同。

		参与者2	
		a	b
参与者1	a	7,7	9,0
	b	0,9	10,10

图 12—11　猎鹿博弈

在没有前博弈沟通的实验中，参与者很少在更好

的均衡点上达成合作，他们几乎都选择了（a,a）。这个结果的产生可能是因为一个参与者认为选 a 至少可以得到收益 7，而如果选 b，得到的收益可能为 0。如果都选 b，情况会更好，但对一个参与者来说是有风险的，因为如果他这样做，他不能保证另一个参与者也做出相同的决定。

当允许单向交流时，在（b, b）点上合作的概率为 53%。这个结果不坏，但真正打动人的是当两个参与者都可以表明他们的意图时：在双向交流的实验中，在（b, b）点上合作的概率高达 90%。

在猎鹿博弈中，双向交流是达到合作最佳均衡的最有效的方法，而相反，在男女之间的博弈中，单向交流是达到合作最佳均衡的最有效的方法。对这个明显的差别的解释是猎鹿博弈中的交流是再次确认每个成员即将做什么。当两个参与者而非一个参与者表明自己的意图是采取 b 时，再次确认的效果就更高。相反，男女之间博弈中每个参与者的信息至少用来向另一方表明他将采取对他最有利的方案。比如，参与者 1 说他打算采取 y 策略，希望在（y, x）上达到均衡。但是因为参与者对均衡有不同的标准，这些信息就对合作起不到任何作用，因为参与者 1 表明他的意图是采取 y 策略，参与者 2 表明他的意图是采取 x 策略。相反，如果只有一个参与者发送短信，那么在前博弈沟通阶段就没有机会争论两者的意图，合作反而更容易达成。

本章总结与练习题

本章总结

采取多种方式进行交流的能力是人类所独有的。尽管长尾黑颚猴在附近有蛇、鹰或者虎的时候有交流的能力，但这远不如人类的复杂。人类的词汇很丰富，手势也多种多样。人类确实可以创造一种语言来表达特定环境下发生的情况，如青少年说“我爸妈在房间”或者一对夫妇在无聊的晚宴上用眼神示意对方是时候回家了。然而，尽管丰富的语言表达能力是交流的必需条件，它不是成功传递信息的充分条件：人类的智力使得人们可以使用复杂的语言来欺骗别人，这就大大减少了我们传送给彼此信息的实质内容。

这一章以**廉价谈话博弈**为背景探索了交流的意义，廉价谈话博弈是一个可以**表明意图**的博弈，在其中信息发送者的选择是**消息**（即不直接影响参与者收益的无成本的行动）。尽管存在胡说均衡，是否存在一个消息有用的均衡取决于信息发送者和接收者之间有没有共同的利益。比如，如果一个投手给接球手示意他将扔一个飙球，接球手就应该相信他，因为投手和接球手都希望把击球手赐出局。然而，如果这两方的利益分歧很大，那么这个示意就没有信息含量。比如，如果接球手私下对击球手说投手将扔一个曲线球，击球手就应该忽略这个信息。因为如果击球手相信了这个消息，接球手会示意投手扔一个曲线球。问题在于接球手想要击球手出局，而击球手想要打中球。本着如此相反的利益，一个参与者说的话就不可信，消息就没有信息含量。

在两种场景中，我们探讨了**前博弈沟通**的作用。在不完全信息博弈中，一条消息使得其中一个参与者知道这个博弈的情况，比如收益。例如，在股票建议博弈中，一个股票分析师对一支股票的走势有了解，他的买入、继续持有或者卖出的建议传达了他所知道而他的客户不知道的信息。这种行为就被称为“**表明信息**”。前博弈沟通在完全信息博弈中也有用，在完全信息博弈中，参与者有共同知识。在这种情况下，消息的作用是传达一个参与者的意图。“表明意图”可以使参与者之间达成合作。在看电影博弈中，菲奥纳和马特如果之前发送短信商量就可以增加他们去同一个电影院的机会。

PDG

练习题

1. 为图 12—12 中的博弈找一个分离精炼贝叶斯纳什均衡。

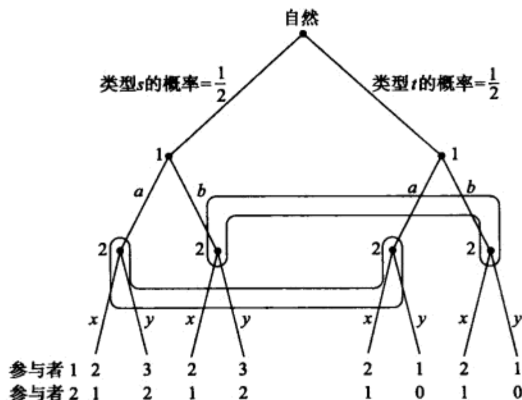


图 12—12

2. 思考在一个廉价谈话博弈中，“自然”选择决定信息发送者的类型，这个类型有四个因素：-1，1，2 和 3，每个出现的概率都为$\frac{1}{4}$。信息发送者知道他的类型且有三种可以选择的消息：x、y 和 z。信息接收者查看发送者的消息，然后有 3 种行动：0、5 和 10。信息发送者的收益是他的类型和信息接收者行动的乘积。信息接收者的收益是发送者的类型与两倍的信息接收者行动的乘积。

a. 求一个分离精炼贝叶斯纳什均衡。

b. 求一个半分离精炼贝叶斯纳什均衡。

3. 思考在一个廉价谈话博弈中，“自然”选择决定信息发送者的类型，这个类型有三个：x，y 和 z，每个出现的概率分别为$\frac{1}{4}$、$\frac{1}{4}$和$\frac{1}{2}$。消息发送者了解他的类型然后选择在 4 条消息中选择：m_1，m_2，m_3 和 m_4。接收者查看发送者的消息然后选择行动：a，b 和 c。收益如表 12—3 所示：

表 12—3

发送者的类型	接收者的行动	发送者的收益	接收者的收益
x	a	3	3
x	b	2	1
x	c	1	2
y	a	4	1
y	b	5	3
y	c	3	4
z	a	3	2
z	b	9	1
z	c	10	0

a. 假设发送者的策略如下：（1）如果类型为 x，那么选择消息 m_1，（2）如果类型为 y 或者 z，那么选择消息 m_2。信息接收者的策略为如下：（1）如果消息为 m_1，那么采取行动 a；（2）如果消息为 m_2，那么采取行动 b；（3）如果消息为 m_3 或 m_4，那么选择行动 a。求解最适合接收者的信念，并证明这个策略组合是一个精炼贝叶斯纳什均衡。

b. 求解一个分离精炼贝叶斯纳什均衡。

4. 回到股票建议博弈，再次思考半分离策略组合，但是让分析师的策略如下：当股票上涨或表现一般时建议买入，股票表现不佳时建议卖出。证明如果 $a+b \leq 2$ 时，那么这个策略组合是部分精炼贝叶斯纳什均衡。

5. 图 12—13 所示是一个廉价谈话博弈，在这个博弈中，发送者有两种类型，表示为 t_1 和 t_2；可以选择两种消息，表示为 m_1 和 m_2。在收到发送者的消息后，接收者有三种可以选择的行动：a, b 和 c。

a. 求一个分离精炼贝叶斯纳什均衡。

b. 求一个混同精炼贝叶斯纳什均衡。

c. 假设发送者是类型 t_1 的概率为 p，是 t_2 的概率为 $1-p$。求存在一组混同精炼贝叶斯纳什均衡 p 的值，且在均衡中接收者采取行动 b。

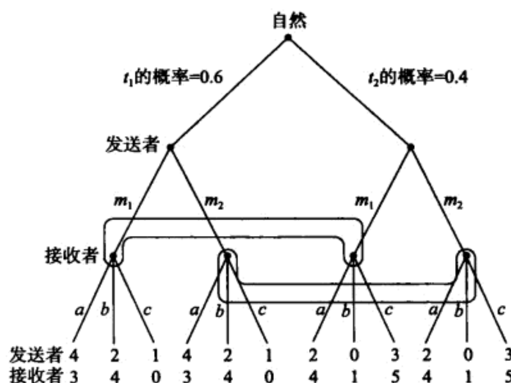


图 12—13

6. 假设格雷丝和莉萨去吃晚餐。格雷丝正要拜访莉萨，她们约好在一家当地的餐馆见面。当莉萨在镇里生活时，有两家不错的餐馆：B 餐馆和 C 餐馆。当然，莉萨知道的消息已过时，但格雷丝知道这些天来，哪家餐馆更好。假设 B 餐馆更好的概率为 $p>\frac{1}{2}$，C 餐馆更好的概率为 $1-p$。“自然”选择决定格雷丝认为更好的餐馆是哪家。格雷丝发送消息给莉萨，内容或者是“让我们去 B 餐馆吃饭”、“让我们去 C 餐馆吃饭”或者是“我不知道哪家餐馆更好”。莉萨收到消息后，格雷丝和莉萨同时决定去哪家餐馆。收益是格雷丝和莉萨想要去同一家餐馆，但是她们更愿意去格雷丝认为更好的那一家。更具体地来说，如果 B 餐馆更好的话，那么她们从采取不同行动中得到的收益如图 12—14a 所示。相反，如果 C 餐馆更好的话，图 12—14b 列出了采取不同行动得到的收益。

		莉萨	
		B餐馆	C餐馆
格雷丝	B餐馆	2,2	0,0
	C餐馆	0,0	1,1

（a） 当 B 餐馆更好时的收益

		莉萨	
		B餐馆	C餐馆
格雷丝	B餐馆	1,1	0,0
	C餐馆	0,0	2,2

（b） 当 C 餐馆更好时的收益

图 12—14

a. 求格雷丝和莉萨经常去更好餐馆的精炼贝叶斯纳什均衡。

b. 求一组混同精炼贝叶斯纳什均衡。

7. 让我们重新看 11.2 小节中的恋爱博弈，但是假设没有礼物。图 12—5 中列出的是这个博弈的扩展形式。这个场景中的私人信息是杰克是否深爱罗斯，并且愿意娶她，同样，罗斯是否深爱杰克，并且愿意嫁给他。只有每个人自己知道自己是否真爱另一方。假设每一个爱对方的概率为 p，这里 $0<p<1$。所以，他们是彼此的另一半的概率，即杰克深爱罗斯，罗斯也深爱杰克的概率为 p^2。

在知道他们的类型后，杰克和罗斯面临一系列的决定：杰克开始向罗斯提出是否可以有婚前性行为。如果他做出这个提议，那么罗斯可以接受，可以拒绝。如果她接受，那么他们有婚前性行为。在这轮提议和行动之后，或者他们结婚（如果他们相爱的话），或者他们不结婚（如果一方不爱另一方）。我们假设并非明确地模拟而是隐含在收益中的婚姻决定和他们是否有婚前性行为无关。杰克的收益取决于他们是否有婚前性行为和是否相爱（所以结婚）。不管他是否爱她，杰克都希望和罗斯做爱，这样杰克从做爱中得到的收益为 $s>0$。如果他和罗斯证明是相爱的且愿意结婚，杰克从结婚中得到的收益 $m>0$。

所以，如果杰克和罗斯有婚前性行为且他们结婚了（因为证明他们是相爱的），那么他得到的收益为 $m+s$。如果他们有婚前性行为，但不能保证最后会结婚，那么他得到的收益只有 s。最后，如果他们既没有婚前性行为也没有结婚，那么他得到的收益为 0。和杰克一样，罗斯从他们结婚中得到收益为 m。对于婚前性行为，她最大的担心就是对方不是未来结婚的那个人。罗斯从他们的婚前性行为和婚姻中得到的收益为 $m+s$，与杰克的相同。然而，她从发生婚前性行为但没有和杰克结婚中（他不爱她或者她不爱他）得到的收益为 $u<0$。最后，她从没有婚前性行为也没有婚姻中得到的收益为 0。证明不存在发生婚前性行为的精炼贝叶斯纳什均衡。

8. 在图 12—16 所示的博弈中，参与者 1 知道他的类型且发送消息，然后参与者 1 和参与者 2 同时决定选择行动。求一个分离精炼贝叶斯纳什均衡。

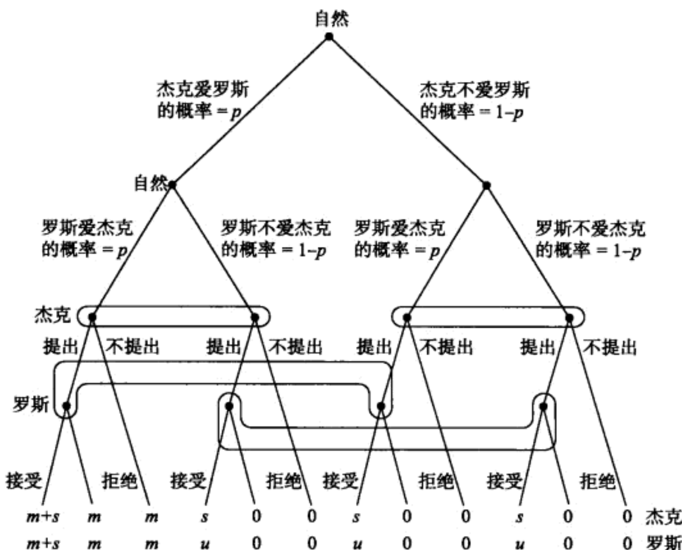


图 12—15　廉价谈话下的恋爱博弈

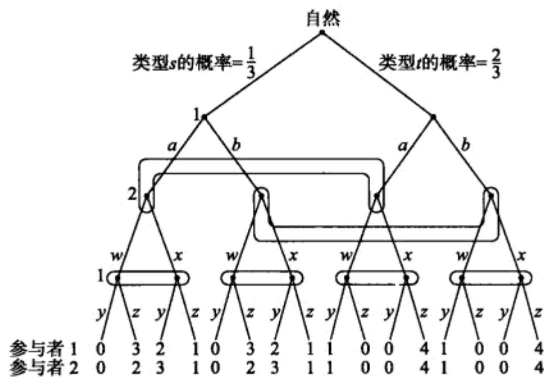


图 12—16　一个博弈

9. 回到 12.3.1 小节中的看电影博弈，但是现在假设菲奥纳和马特在前博弈沟通中都有三种消息：“让我们去看肥皂剧吧”、“让我们去看动作片吧”和“聚会去吧”。推导 a, b, y 和 z 的值使得下面的策略组合是一个子博弈精炼纳什均衡。

■ **马特的策略：**

发送短信“让我们去看肥皂剧吧”的概率为 a，“让我们去看动作片吧”的概率为 b，“聚会去吧”的概率为 $1-a-b$。

如果（1）他们都发送“让我们去看肥皂剧吧”或（2）一个发短信“让我们去看肥皂剧吧”，另一个发短信“聚会去吧”，那么就去看肥皂剧。

如果（1）他们都发送“让我们去看动作片吧”或（2）一个发短信“让我们去看动作片吧”，另一个发短信“聚会去吧”，那么就去看动作片。

对任何其他的信息组合，去看肥皂剧的概率为 $\frac{1}{3}$。

■ **菲奥纳的策略：**

发送短信“让我们去看肥皂剧吧”的概率为 y，“让我们去看动作片吧”的概率为 z，“聚会去吧”的概率为 $1-y-z$。

如果（1）他们都发送“让我们去看肥皂剧吧”或（2）一个发短信“让我们去看肥皂剧吧”，另一个发短信“聚会去吧”，那么就去看肥皂剧。

如果（1）他们都发送“让我们去看动作片吧”或（2）一个发短信“让我们去看动作片吧”，另一个发短信“聚会去吧”，那么就去看动作片。

对任何其他的信息组合，去看肥皂剧的概率为 $\frac{3}{4}$。

10. 回到 12.3.1 小节中的看电影博弈，但现在假设不是有一轮，而是有两轮前博弈沟通。求一个子博弈精炼纳什均衡。

第五部分 重复博弈

第 13 章

参与到永远：非固定居所参与者的重复互动博弈

学习目标

13.1 创建一个重复博弈
13.2 堑壕战：有限次视野
13.3 堑壕战：无限次视野
13.4 针对重复“囚徒困境”的实验证据

本章导读

有这样一些人，他们既没有开枪杀过人也没有听到过那些流血的伤者们痛苦的尖叫、呻吟。只有这样的人才更想复仇，他们的内心更加悲伤。战争是地狱。

——威廉·谢尔曼将军

战争让人类的道德约束力量荡然无存。现实社会，杀害一个人会被定为谋杀罪，而战争期间杀人是士兵的义务。第一次世界大战中人类的死伤数量是骇人听闻的，而这一数字也成了日后战争进行比较的标准。这种杀戮大多发生在法国及其他战区构筑的堑壕里。协约国和德国展开的是一场持久战，因为他们都会定期从土筑的防御工事中向对方发起进攻。

在这场流血的战争中人性并没有完全泯灭，期间还会出现短暂的和平。因为对面堑壕里的士兵时而会达成停战协议。他们会定期发起攻击而非突然袭击以使对方有足够的时间做好防御。他们约好不在吃饭时相互进攻，也不炮击对方的攻击线。

场景：（营地）上午 8 点到 9 点是士兵处理个人事务的时间，而且旗帜标识的特定区域被双方的狙击手认为是非射击区。

在前线一两天之后，琼斯上校发现了德国人的规律：他们在目标选择、射击时间以及轰炸间隔等方面都是有规律的。他能够在一分钟之内预测下一发炮弹的降落位置。他的推算是非常准确的，而且作为文职人员，他能够冒一个看似很大的风险，他能够预知在他到达被炮轰区域之前炮轰会停止。

我正与一位战友喝茶时，听到了喧闹声，出去一看才发现我们的军队与德方军队站在各自的防线前。突然一颗流弹飞来，幸好没有造成伤害。双方士兵很自然地卧倒在地，我方士兵开始咒骂德国人。突然之间，一个勇敢的德国士兵站在防线前，对我们

喊道："非常抱歉，我们并不希望有人受伤。这并不是我们的过错，而是那些该死的普鲁士炮兵"。

当杀戮成为例行公事的时候，这些士兵就能达成协议，停止相互残杀。这种形式的合作是如何取得并得到维系的呢？正如你所期望的那样，博弈论能够为我们解决这些疑惑提供一线曙光。最后，我们将根据堑壕里士兵所面临的情形创建一种策略式博弈。



第一次世界大战堑壕里的士兵

为简化分析，假设这场战争只有两个参与者，协约国士兵（看成一个团体）和德国士兵（看成一个团体），每个参与者有两个策略，即士兵可以试着杀死敌人（击中敌人）或者不杀（未击中）。当然，这种单纯的假设士兵不愿杀死敌人不足以导致这种友好行为的发生。但是那只是一个假设，而这种假设在许多士兵试着杀戮的事例中是不成立的。确切地说，假设每个士兵都重视杀敌所产生的收益，但更看重保全自身所产生的更高的收益。士兵的收益可以具体的表示成：

4+2×（敌方士兵被杀数量）−4×（我方士兵被杀数量）

我们注意到越多敌人被杀，士兵所获得的收益越高，而越多自己的战友被杀所获得的收益越低。此外，就宿命而言，一对一的交换会令士兵的境况更糟。因为每牺牲一个自己的战友他减少 4 个单位的收益，而每杀死一个敌人他增加两个单位的收益。因此，其收益净减少两个单位。

接着，假设开始射击产生两种结果，杀死一个敌人或者未击中任何人，这一说明会导致图 13—1 所示的策略式博弈。

德国士兵

		击中	不击中
协约国士兵	击中	2,2	6,0
	不击中	0,6	4,4

图 13—1　堑壕战博弈

这里有一个唯一的纳什均衡，即双方都试图杀死对方。这一解释非常令人信服，因为选择去杀人为占优策略（实际上这次博弈仅仅是"囚徒困境"）。不管敌方做何种选择，士兵开枪射击是比较明智的。他的行为不会影响战友的牺牲数量（这是由敌人决定的），他只能决定杀死多少敌人。对自己来说，消灭的敌人越多越好，因此开枪杀敌是最佳策略。然而，如果双方能够合作并达成协议，相约不试图杀害对方，那么他们的境况都会好转。因为这一决策会让每个参与者的收益从 2 增加到 4。

很显然我们已经无法解释为什么在第一次世界大战的堑壕里会出现周期性的和平了。是我们哪里弄错了吗？难道是我们错判了士兵的喜好？亦或是我们对目标的假设不合理？为了寻求突破，我们可以这样问自己：为什么这种和平行为会出现在第一次世界大战的堑壕中，而在大多数其他战争里却很少出现呢？堑壕战的一个独特特点是双方的交战是周期性的。他们在战斗最激烈的时候决定是否杀敌不是一次性的，他们要日复一日的面对相同的情景并做

PDG

出同样的决定。由于这种战争所具有的特性，士兵必须一直周期性地面对对方。

这种策略性互动的重复是解决为什么在第一次世界大战的堑壕里会出现和平行为的关键，也是本章节所要探讨的。在许多策略性场景中（不仅仅局限于堑壕），重复遭遇能够维持合作，在重复博弈中通过合作每个人获得的收益要高于一次性博弈产生的收益。在 13.1 节，我们会正式创建一场周期性交战的博弈。接着在 13.2 节和 13.3 节中，我们分析这场堑壕战博弈，首先，假设交战士兵的数量是固定的（比如说 7 个），接着假设他们的交战状态能够无限持续下去。13.4 节提供了一些相关的实验性证据来说明实际上还有很多人类的行为是我们所不理解的。第 14 章接着列举的一些例子，即人类重复地影响着彼此的行为，扩宽了我们的视野。

13.1 创建一个重复博弈

所谓**重复博弈**，就是指所有参与者周而复始地面对相同的遭遇——**阶段博弈**（stage game）的一种情形。阶段博弈是构成重复博弈的组件，或者可以构成有几个固定步骤的其他博弈。例如，图 13—1 中的博弈就是阶段博弈，它将最终演变成一场堑壕战。从阶段博弈到重复博弈的演变过程中，我们有必要重新定义一下游戏规则和收益机制。因为策略完全是为博弈规则制定的，所以如果一个参与者被期望有多次遭遇而不止一次，那么这一套行之有效的规则将有不同的表现。至于收益，我们会很自然地认为参与者不单考虑当前处境所产生的收益，也会考虑所有未来可能的遭遇。

假设协约国和德国士兵预计会按照图 13—1 所示的方式交战 T 次。他们每天交战一次共持续 T 天。图 13—1 所示的博弈中，击中是一个策略，但是在重复博弈中这一策略更加复杂。阶段博弈中“击中或者不击中”是一个策略，而在重复博弈中这一策略被定义为一种行为。在重复博弈中策略的概念等同于其他任何形式的博弈。对于一个参与者来说，策略仅适用于根据每个信息所设定的行为。因此，一个策略如何取决于他做出决定时所掌握的信息。

我们设想两种不同的信息模式，令 $T=2$，那么阶段博弈仅出现两次。这也就意味着第一天，协约国士兵和德国士兵必须同时做出是否击中的决定，第二天也一样。图 13—2 是一个扩展形式的博弈，第一天所发生的事情在第二天并不为参与者知晓（现在我们先忽略收益，稍后会做出解释）。我们注意到每个参与者都知道他第一天所做的事情（否则，这场博弈不满足完美记忆的要求），但是当他决定第二天该采取何种行为时，他并不知悉其他参与者第一天的行为。协约国的士兵有三种信息模式：（1）第一天获取的原始信息；（2）与第一天选择击中有关的信息；（3）与第一天选择不击中敌人有关的信息。因此，一种策略包含 3 种行为，且一种策略包含 8 种方式，这 8 种方式都与 3 种信息模式下选择是否击中有关。

将这一博弈与图 13—3 所示的扩展形式的博弈对比会发现它们的主干几乎是一样的，所不同的仅仅是消息模式的结构。现在假设第一天所做的行为会影响两个参与者第二天的选择。每一个玩家都有 5 种信息模式，一种与第一天有关而剩下的 4 种与第二天相关。参与者可能知道第一天所发生的 4 种情形，因此有 4 种信息模式与第二天相关：（1）双方都选择击中；（2）协约国士兵选择击中而德国士兵选择不击中敌人；（3）协约国士兵选择不击中敌人而德国士兵选择击中；（4）双方都选择不击中敌人。参与者的每一个策略都对应 5 种行为，每种行为对应 5 种信息模式，共有 32（$=2^5$）种可行策略。

本章将重点探讨一个历史（即参与者过去的选择）是共有知识的案例，正如图 13—3 中的博弈所反应的。当过去的选择成为一种共有知识且每个阶段博弈进行 3 次，此时就会形成 21 种信息模式：与第一天相关的 1 种信息，与第二天相关的 4 种信息，与第三天相关的 16 种信息。第三天的信息总结了前两天所发生的事情。这可能表现为两个参与者在第一天都选择不击中敌人，第二天协约国士兵选择击中而德国士兵选择不击中敌人。前两天共产生了 16 种可能事件，因此第三天共有 16 种信息模式。这样的话，为期三个阶段的重复博弈所形成的策略是一个包含“21 元组”可能行为的组合。

概括地讲，一个策略决定每一阶段的行为，以及前面阶段的（参与者所熟知的）可能性。维持 T 个阶段的博弈，每一个阶段博弈中每个参与者都有 m 种可能的行为。这样就会在第一阶段形成一个初始的信息，与第二阶段相关的信息为 m^2，与第三个阶段相关的信息共有

m^4……到阶段 T，共有 $m^{2(T-1)}$ 种信息。从而，一个策略就是这么多种信息模式下可能行为的组合，即：

$$1+m^2+m^4+\cdots+m^{2(T-1)}$$

对于一场不确定有多少阶段的博弈（$T=\infty$），由于与此相关的信息模式的数量是不确定的，此策略所包含的可能行为的数量也是不确定的。一个策略可能超乎想象的复杂，在此我们暂不做讨论。实际上，我们可以看出一个简单的策略就可以重新解释关于堑壕战的疑惑。

重复博弈的另一个组成要素便是收益。就像阶段博弈中的策略在重复博弈中表现为一种行动，阶段博弈中的收益在重复博弈中仅表现为一个阶段的收益。重复博弈中参与者的收益受到每一个独立阶段所获收益的影响。假如 $T=5$，那么共有以下几种可能：

（不击中，不击中），（击中，不击中），（不击中，不击中），（不击中，击中），（击中，击中）

在每一对有序的行为组合中首次的行动都是因协约国一方而生，从图 13—1 可知，协约国士兵每一个独立阶段的收益为 {4,6,4,0,2}。

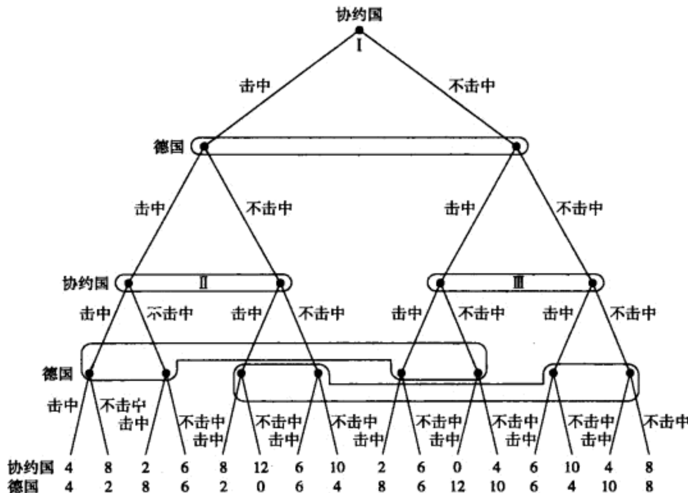


图 13—2　过去的行为不被知晓的情况下的第二阶段的堑壕战

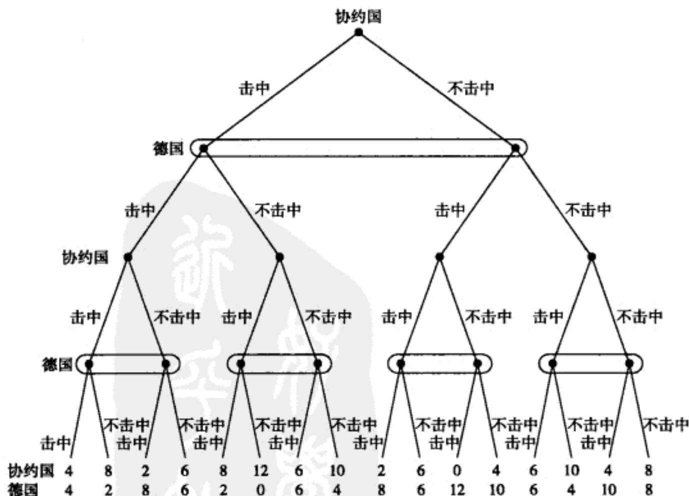


图 13—3　过去的行为是共有知识情况下的第二阶段的堑壕战

假设在重复博弈中参与者会选择能够产生最高收益的策略作为其最终策略。但是一种策略所产生的收益高于另一种策略所产生的收益又意味着什么呢？对比两组不同的收益 {4,6,4,0,2} 与 {6,6,6,0,2}，很显然后者会更受青睐，因为在阶段 1 和阶段 3 它创造了更高的收益而在其他阶段也产生了同样的收益。假设用一组收益 {6,4,4,2,6} 代替 {6,6,6,0,2} 这组收益，此时与 {4,6,4,0,2} 这组收益相比，我们不能很容易地发现哪个更好，因为在 1,4,5 阶段，{6,4,4,2,6} 这组收益更具吸引力，而在阶段 2，{4,6,4,0,2} 这组收益更好。

我们需要一个标准以进行收益比较。通常为了得到这一标准都是把所得收益转化成简单的数字（我们称之为总计数据），接着假设一个参与者所选择的策略是为了使总计数据最大化。为了得到这一总计数据，我们会很自然地想到各个单一阶段的收益之和。这样做不只是因为简单，更重要的当单一阶段的收益增加时总计数据会随之增加。对于上述的三组收益，表 13—1 给出了其总计数据。如果使用我们刚刚提到的这些数据，协约国士兵会按照 $C > B > A$ 的顺序排序。（图 13—2 及图 13—3 用到了单一阶段的收益之和）。

表13—1　收益总计

	阶段收益	总计数据
A	4,6,4,0,2	16
B	6,6,6,0,2	20
C	6,4,4,2,6	22

很多时候，使用单一阶段的收益总计并不能达到满意的效果，因为每一阶段的收益被认为是同等重要的。例如，如果你仅希望在夏季的工作中让自己的收入达到最大值，那么你将不在乎是每周末拿到薪水还是到夏季末拿到薪酬。但是如果跟大多数人一样，你更希望每周拿一次薪水，即使你打算到夏季末再消费这笔钱，因为你可以把较早到手的钱存入银行以获得利息。总之，每个人都希望尽早拿到薪水（得到收益）。

基于这种想法，我们会更看重及早到手的收益。我们考虑另一种带权重的收益之和，且越往后一阶段数据的权重越小，而不是单纯地把每一个独立阶段的收益加总。我们用 u_t 表示阶段 t 所得的收益，那么总计收益表示为：

$$w_1u_1+w_2u_2+w_3u_3+w_4u_4+\cdots+w_Tu_T$$

其中：

$$w_1>w_2>\cdots>w_T>0$$

经济学家喜欢用下列形式表示权重 w_t：

$$w_t=\delta^{t-1}，其中\ 0\leqslant\delta\leqslant1$$

总之，阶段 t 上的数据的权重等同于分数的"t-1"次方。我们都知道一个数字乘以分数就会变小，因此权重一定会减小（除非 δ 为 0 或 1，此时它们分别恒等于 0 或 1）如果 δ=0.6，那么 w_1=1，w_2=0.6，w_3=0.36，w_4=0.216，…

当权重以这种形式表示时，附着要素之和被认为是单一阶段收益的现值，表示为 $u_1+\delta u_2+\delta^2u_3+\delta^3u_4+\cdots+\delta^{T-1}u_T$。其中 δ 表示折现系数，因为它是用来衡量未来折扣情况的要素。收益的现值是指打折后的单一阶段收益的和。不熟悉现值的同学可以翻阅本章附录 1，其中介绍了一些关于现值的特性，将对以后的应用有所帮助。

13.1 理解测验

假设δ=0.8，确定表13—2所示的三组收益中哪个产生的现值最高。

表13—2　不同的收益组合

阶段	组合A	组合B	组合C
1	15	25	5
2	15	15	10
3	15	10	20
4	15	5	30

13.2 堑壕战：有限次视野

杀戮是被禁止的，因此所有的杀人犯都会受到惩罚，除非他们大肆杀戮是因为听到了冲锋号。

——伏尔泰

随着重复博弈的建立，我们现在已经能够解释为什么在法国的堑壕里会出现短暂的和平了。首先假设协约国士兵和德国士兵进行了两次对决（T=2），并且每个参与者都试图使单一阶段的收益之和最大化（δ=1）。此时博弈情况如图 13—4 所示（回顾前面的章节，我们假设过去的选择是共同知识）。

这场博弈其实只是第 8 章中我们所探讨的博弈的扩展形式。回顾前面的内容我们选择的解决方案是一

个子博弈精炼纳什均衡，而这一均衡可以用逆向归纳法解决。从图 13—3 中，我们可以看到共有 5 个子博弈：这场博弈自身及第二阶段的 4 个子博弈。逆向归纳法让我们能够解决纳什均衡的 4 个子博弈。

我们从探讨双方在第一阶段都选择击中策略的子博弈开始。图 13—4 给出了这一策略的表现形式，且我们很容易证实它有唯一的纳什均衡（击中，击中）。因此，如果双方在第一阶段都选择了击中，子博弈精炼纳什均衡会令双方在第二阶段做出同样的选择。

德国士兵

		击中	不击中
协约国士兵	击中	4,4	8,2
	不击中	2,8	6,6

图 13—4　阶段一（击中，击中）之后堑壕战博弈第二阶段的子博弈

现在思考这样一个子博弈：第一阶段，协约国士兵选择开枪杀敌而德国士兵不击中敌人。图 13—5 描述了这一策略形式。这里，（击中，击中）还是唯一的一个纳什均衡，我将留给你自己去证实：（击中，击中）在其他两个阶段的子博弈中也是唯一的一个纳什均衡。

德国士兵

		击中	不击中
协约国士兵	击中	8,2	12,0
	不击中	6,6	10,4

图 13—5　阶段一（击中，不击中）之后堑壕战博弈第二阶段的子博弈

作为逆向归纳法的一部分，4 个子博弈中的每一个都被相关的纳什均衡的收益所替代。按照这一步骤发展下去会导致图 13—6 所示的结果，它清楚明了地表明所描述的博弈有唯一的一个纳什均衡。而两个参与者都选择击中这一策略是出人意料的。如果把所有的分析连在一起，我们会发现图 13—3 中第二阶段的堑壕战博弈有唯一的子博弈精炼纳什均衡：就其 5 种信息模式而言，每一个模式都令双方将选择击中作为策略。而这场博弈的一系列结果使他们在阶段 1 和阶段 2 都选择击中。

那么，上次应用是最大的错误，因为我们始终没有接近这一目标：产生和平行为并作为均衡条件的一部分。我们认为更长时期的交战对我们的分析毫无帮助。不管历经 10 个阶段还是 100 个、1 000 个阶段甚至 100 万个阶段，都只有双方都试图击中对方这一种均衡行为。

假设一场博弈要经历 T 个阶段而此时是最后阶段，总收益就是各个独立阶段的收益之和（尽管我们使用添加权重求和这一方法甚至会引发非议）。令 A^{T-1} 和 G^{T-1} 分别表示第一个 T-1 阶段协约国和德国方面单一阶段的收益之和。可以看到，T 阶段面对的子博弈如图 13—7 所示。我们所做的仅是用图 13—1 表示这场博弈，并且将 A^{T-1} 加到每个协约国士兵的收益上，将 G^{T-1} 加到每个德国士兵的收益上。当然，如果在一次击中博弈中，“击中”压倒“不击中”占主导地位，那么就算我们给每个收益加上一个常数，结果还是一样的（要让自己明白这一事实）。最后阶段，很显然对双方来说，击中对方是最佳策略。另一个考虑这一结果的方法就是注意经过第一个 T-1 阶段，协约国一方已经意识到 A^{T-1} 这一收益，他们是这样计划的也是照此行动的。到了阶段 T，他们应该尽最大努力杀死尽可能多的德国士兵以求得最高收益。

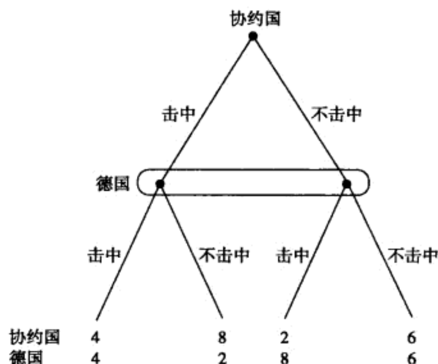


图 13—6　逆向归纳后的两阶段堑壕战的第一阶段

德国士兵

		击中	不击中
协约国士兵	击中	$A^{T-1}+2, G^{T-1}+2$	$A^{T-1}+6, G^{T-1}$
	不击中	$A^{T-1}, G^{T-1}+6$	$A^{T-1}+4, G^{T-1}+4$

图 13—7

现在我们回到T-1阶段，姑且不论T-1阶段发生了什么，正如我们之前讨论的，最后阶段双方都会选择击中。假设A^{T-2}及G^{T-2}是第一个T-2阶段单一阶段的收益之和。图13—8展示了T-1阶段士兵面对的博弈。我

们再次只是在图13—1的收益上加上一个常数，现在是$A^{T-2}+2$或$G^{T-2}+2$。因此，依然是击中压倒不击中占据优势。所以在倒数第二个阶段双方依然会选择击中作为策略。

德国士兵

协约国士兵		击中	不击中
	击中	$A^{T-2}+4,G^{T-1}+4$	$A^{T-1}+8,G^{T-1}+2$
	不击中	$A^{T-1}+2,G^{T-1}+8$	$A^{T-1}+6,G^{T-1}+6$

图 13—8

这一讨论可以说明双方在 T-2，T-3 阶段直到第一阶段都会选择击中作为策略。考虑一下为什么会这样。在阶段 T，所得收益超越了第一个 T-1 阶段且不为士兵 T 阶段的行为所影响。最明智的行为意味着选择最大化当前收益的策略。正如图 13—1 所示的博弈，击中显然是最佳策略。到了阶段 T-1，前几个阶段的收益也是固定的且不受阶段 T-1 策略的影响。那么不管阶段 T-1 发生什么，从阶段 T 开始，未来期望收益为 2（因为我们认为到最后阶段双方都会选择击中）。阶段 T-1 的情形仍然像一次性博弈：T-1 阶段的行为只会影响 T-1 阶段的收益，且 T-2 阶段的行为只会影响 T-2 阶段的收益，总之过去的事情是无法改变的，将来要发生什么也是不受其他阶段影响的，因为双方总是试图击中对方。也就是说，T-2 阶段的情形依然像一次性博弈。这种逻辑可以被持续应用到未来任一阶段以表明在任一阶段双方都会击中。

下面的消息比听起来更糟，因为这一逻辑既可以运用到重复的堑壕战博弈中，也可以运用到任何有特定阶段的重复博弈中，且这种博弈的阶段博弈有唯一的纳什均衡。对于阶段博弈来说，有一个纳什均衡的重要性在于它能够明确最后一个阶段参与者如何表现。因此可以得出这一结论，即倒数第二阶段的行为不会影响到最后一个阶段，并且这一逻辑可以重复地应用到任何阶段直到初始阶段。


点拨

如果阶段博弈有唯一的纳什均衡，对于一个有限次重复博弈来说，重复阶段博弈中的纳什均衡是完善纳什均衡路线的唯一的子博弈精炼纳什均衡。没有永远的合作。

13.2 理解测验

假设参与者1和参与者2参与了3次如图13—9所示的阶段博弈，所有三阶段他们所得的收益就是各个单阶段的收益之和。子博弈精炼纳什均衡暗示着怎样的行为顺序呢？

参与者 2

参与者 1		低	中等	高
	低	2,2	6,1	3,3
	中等	1,6	5,5	2,1
	高	3,3	1,2	0,0

图 13—9

13.3 堑壕战：无限次视野

我们总是在重复我们做过的事。

——亚里士多德

有限次重复的堑壕战博弈的一个重要特征在于，参与者确切地知道他们的遭遇战何时停止。同时，这一特征也是分析的关键，因为我们要讨论何时是一场博弈的最后阶段。参与者知道这是最后一次了，所以他们把自己面对的博弈看成是一次性博弈。尽管“博弈的次数是有限的”这一假设或许是合理的（毕竟我们的寿命有限），然而在大多数情况下“参与者确切地知道博弈何时停止”这一假设依然是不合理的。

如果一场博弈不具备“信息透明”这一特性就意味着这场博弈所持续的阶段是**不确定视野**（indefinite horizon）。不确定视野就是说这场博弈有持续下去的可能。例如，假设每一阶段，参与者在第二天又与对方交手的概率为 p，因此当前的交手是他们最后一次交手的概率为 $1-p$。有一点是很重要的，参与者在决定如何做时，他们不确定第二天是否会与对方再次交手（他们的关系有持续下去的可能）。当 p=1 时，博弈具有**无限次视野**（infinite horizon），参与者的交手也肯定会持续下去。

正如本章附录 2 所解释的，这场博弈所持续的阶段是不确定的（$0<p<1$）还是无限次的（p=1）将导致不同的结果。我们假设这场博弈是无限次的，但是也要记住“参与者从来都不知道当前阶段是他们相互影响的最

后阶段”是这一假设成立的决定性条件。

在无限次重复博弈中，用当前评价标准去评价单一阶段的收益并且假设折现系数 δ 小于 1，这两个条件是很重要的。为了寻找原因，假设 $\delta=1$ 且每个阶段都会产生 5 个单位的收益，那么总收益就是无限个 5 的加总。处理如此庞大的数字是一件令人厌恶的工作。就像尤达（Yoda）所说的“我不想计算这些数字”。但是，如果 $0<\delta<1$，则相同的收益产生的现值为：

$$5+\delta 5+\delta^2 5+\delta^3 5+\cdots \text{或} \sum_{t=1}^{\infty} \delta^{t-1}5$$

这个式子也可以表示成 5 /（$1-\delta$），这种表达是一个很好的，容易处理的有限的数字（证明参照本章附录 2）。例如，当 $\delta=0.75$ 时，收益为 20（=5/（1−0.75））。

无限次重复博弈的策略规定了每个历史时期每个阶段参与者的行为。不幸的是，因为不存在最后的子博弈去开始这个过程，逆向归纳法并不适用于无限次视野。同时，尽管我们缺少一个快速想出子博弈精炼纳什均衡的方法，但是依然可以提出一些候选策略，看是否是纳什均衡。

第 9 章，如果对每个子博弈而言，子策略就是一个纳什均衡，那么这个策略就被定义为子博弈精炼纳什均衡。此时我们回想起这样的内容，一个子策略的子博弈是参与者策略的一部分，这个子博弈仅仅决定了由子博弈中的信息模式所规定的行为。而这个定义可以被完善以适应重复博弈的语境。

定义 13.1

重复博弈中子博弈精炼纳什均衡（Subgame Perfect Nash Equilibrium for a Repeated Game）：对一个重复博弈而言，当且仅当这一条件成立时，即每个历史时期的每一阶段对于参与者来说策略所规定的行为都是最佳选择，这个策略就是子博弈精炼纳什均衡，给定（1）其他参与者在当前阶段按照他们自己的策略行事；（2）所有参与者（包括替补参与者）将来按照他们的策略行事。换句话说，假如其他参与者按照他们的策略行事且第一个参与者将来按照她的策略行事，此时那个特定的参与者的策略就能指定最佳行为。

精明的参与者或许会怀疑我试图找到一个快速的解决方法。这里需要注意一下，我们仅需要一个忠于自己策略的参与者，就像我们认为在一些背叛行为发生后他会遵从自己的策略。但是，难道我们不需要确认一下发生一系列背叛行为并不是明智的吗？难道我们没有必要证明这一策略优于其他策略吗？决对有必要，本章附录中，我们讨论了这个问题，如果背叛一次（任何历史时期的任何阶段）是不明智的，那么多次背叛策略（甚至是无限次的）同样是不明智的。

考虑一下下面这个异常简单的策略：任何阶段，任何时候，参与者都选择击中。尽管过去这一策略并没有令每一个参与者为他所做的选择负责，但是请记住，这仅仅是因为过去对每一个参与者而言，让自己的行为可能发生而非必须做出某种行为是有可行性的。接着让我们证明一下如果协约国和德国士兵都选择这一策略，那么对无限次重复的堑壕战博弈而言，它将是一个子博弈精炼纳什均衡。

在这个简单的博弈中，不管处于哪一阶段，哪个历史时期，两个参与者的策略都要求他们做出同样的行为：击中。这一选择使参与者获得的当前收益为

$$2+\delta 2+\delta^2 2+\delta^3 2+\cdots=\frac{2}{1-\delta} \qquad (13.1)$$

这个参与者也希望敌人开枪（根据他的策略），并且希望未来双方都能开枪（据双方策略）。从而，当前阶段及所假设的每一阶段参与者都希望得到两个单位的收益。为了对比，可以假设参与者不击中敌人，此时收益为：

$$0+\delta 2+\delta^2 2+\delta^3 2+\cdots=\delta(2+\delta 2+\delta^2 2+\cdots)=\delta\left(\frac{2}{1-\delta}\right) \qquad (13.2)$$

根据这一策略，参与者今天将得到零收益（参与者希望敌人击中）而未来得到两个单位的收益（据他们的策略，出于鲁莽双方都会选择击中）。很显然，当 $\delta<1$ 时，$\frac{2}{1-\delta}>\delta\left(\frac{2}{1-\delta}\right)$ 成立，所以击中产生的收益较高。还有另一点可以解释为什么开枪射击这一策略更受参与者青睐，那就是当前阶段此策略产生较高的收益（2 对 0），且未来所得收益与另一策略产生的收益相同。

刚刚我们证明了只要双方一直交战，每一阶段双方都击中就是一个子博弈精炼纳什均衡。但是，第一次世界大战期间有些堑壕里产生的和平合作性均衡是怎么回事呢？思考一下下面的这对策略组合：

- 阶段 1：选择不击中。

■ 阶段 t（$t \geqslant 2$）：

如果双方军队过去都选择不击中，则选择不击中；

其他任何历史时期选择击中。

如果双方都采用这一策略，那么他们将选择合作的方式以不击中对方开始这场博弈。只要双方一直遵守合作原则，没有谁先去击中对方，这一规则会永远得以存续。但是只要有人违背游戏规则（一方首先击中），那么之后双方都会放弃合作选择击中。这就是我们所说的 **"冷酷策略"**（grim-trigger strategy）。任何违背游戏规则的行为都会招致惩罚，那就是双方在未来确定的阶段都会击中。

如果双方都采用"冷酷策略"将会产生大家所渴望的休战，即每个阶段参与者都不会击中对方。我们所要做的就是证明这个策略是一个均衡，更具体的证明正是基于这一策略参与者在任何历史时期的任何阶段都做出了最明智的选择。从表面上看，这确实很难。因为这场博弈经历了无限个阶段，无限个历史时期。虽然如此，我们姑且不论它令人畏惧的外表，只需考虑两个案例。

首先，设想这样一个阶段，在此期间没有人选择击中。这个阶段可以是第一阶段也可以是之后的任何阶段。不过之前双方都选择不击中对方。如果一个军队的策略让他不击中对方，那么其期望收益为：

$$4+\delta 4+\delta^2 4+\delta^3 4+\cdots=\frac{4}{1-\delta} \quad (13.3)$$

因为他也不希望对方击中自己，并且希望未来双方都不击中彼此。为了证明后一主张的正确性，记住这一点，每个参与者都希望所有的参与者未来按照他们自己的策略行事。如果双方当前阶段都选择不击中对方，那么根据他们的策略，下一阶段他们也不会击中对方。因为之前的任何阶段没有人击中，这个推理也适用于之后的任何阶段。

这一策略产生的收益在当前阶段大于等于其他策略产生的收益时才能达到均衡。唯一的选择就是击中，并且击中所产生的收益为：

$$6+\delta 2+\delta^2 2+\delta^3 2+\cdots=6+\delta\left(\frac{2}{1-\delta}\right) \quad (13.4)$$

所以，对方失误而自己击中时能获得较高的当前收益 6，但是这个策略是以得到应有的惩罚为代价的：据双方的策略，他们在确定的阶段通过击中的方式回应那些同样击中的参与者，这样获得较低的收益 2。

满足均衡条件需要（13.3）大于等于（13.4），即

$$\frac{4}{1-\delta} \geqslant 6+\delta\left(\frac{2}{1-\delta}\right) \quad (13.5)$$

此时每一方参与者都愿意选择"不击中"作为自己的策略。不等式（13.5）必须满足，这一对策略才成为一个子博弈精炼纳什均衡。

回想起这样一个定律：**一个策略必须为每一历史时期每一阶段的行为提供最准确的指示。** 其实我们只是处理了其中的一部分，还有相当大一部分等待我们去处理，即期间有些参与者击中的情况。据他们的策略，一个军队在这样的阶段会选择击中。假设另一方军队也选择击中且双方在确定的阶段将继续这一选择，此时的收益如式子（13.1）所示。唯一的选择就是"不击中"，所获得的收益如式子（13.2）所示。均衡条件要求式子（13.1）大于等于式子（13.2），我们已证明了这一条件的正确性。

我们已经评价了每一历史时期每一阶段这一策略的最优性。现在我想说的是这一策略并不是长久的最佳策略，为什么呢？我们的结论是：当且仅当不等式（13.5）成立时，"冷酷策略"才是一个子博弈精炼纳什均衡。在这一条件下重新排列式子的次序更有利于揭示问题的真相，所以我们重新处理一下不等式（13.5）：

$$(1-\delta)\times\left(\frac{4}{1-\delta}\right) \geqslant (1-\delta)\times\left[6+\delta\left(\frac{2}{1-\delta}\right)\right]$$

$$4 \geqslant (1-\delta)6+\delta 2 \Rightarrow 4\delta \geqslant 2 \Rightarrow \delta \geqslant \frac{1}{2}$$

从上述式子可以看到，如果 $\delta \geqslant \frac{1}{2}$，这个"冷酷策略"就能满足均衡条件，而如果 $\delta < \frac{1}{2}$，则不能达到均衡。到此为止，第一次世界大战的堑壕中会出现短暂的和平这一问题得到了圆满的解决！

满足条件 $\delta \geqslant \frac{1}{2}$ 是维持合作关系的一般原则。我们设想双方一直保持一致，不击中对方。图 13—10 描述了单一阶段继续不击中对方或者击中对方这两种情况下所产生的收益。如前所述，如果一方选择不击中对方，并且希望继续维系合作关系，他所获得的当前收益及未来收益都为 4。相反，如果有一方击中对方，双方都会卷入残酷的战争，此时他将获得 6 个单位的当前收益，但是未来收益仅为 2。因为每个军队都面对着一场"交易"。如果他击中对方，当前收益从 4 升到 6，正如图 13—10 所示，这是一个逐渐变化的领域，被命名为"当前收益"。但是收益的增加也伴随的成本的付出。双方的状态从"停止交火"转变为"愿意交火"，因此未来收益从 4 降到 2。而"未来收益的丢失"也在图 13—10 中描述了。为

了达到均衡，由在堑壕战中“背叛协议”导致的未来收益的丢失必须高于当前收益。当一个参与者的未来福利足够重要时，“长期的收益丢失超越了短期所得的收益”这一原则就是正确的。因为 δ 的值越大，表示给未来收益更多的权重，δ 必须足够大以使合作是最优策略。当然，我们的条件非常明确：为了使“冷酷策略”达到均衡，δ 必须大于等于 $\frac{1}{2}$。

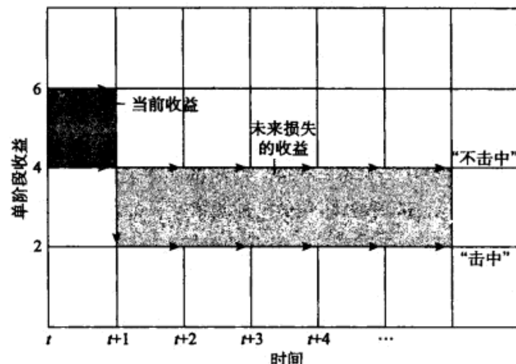


图 13—10　背叛协议后所得收益与损失的收益之间的比较

在我们的分析中最重要的是奖惩方案。一方面，如果协约国士兵维持休战协定，德军方面会在未来报答他们的这一行为，未来作战时试着不击中他们。另一方面，如果协约国一方违背了休战协定，那么德军方面也会做出回应，惩罚这一行为。同时，德军方面维持或违背休战协定也会得到相同的结果。从而，我们得出这一结论，和平的维持不是靠和平主义或者对同伴的爱，而是出于狭隘的个人利益：为了减少造成未来互相残杀的概率，每一方都在努力维持和平。

尽管我们只是分析了一场具体的博弈（堑壕战），一套具体的策略（冷酷策略），与图 13—10 所示的“交易”是大体相同的，维系合作关系的方案是普遍通用的。考虑到任何博弈中，阶段博弈的纳什均衡并不都是最佳的，对参与者来说，这里有另外的一系列行动作为选择以使所有的参与者保持休战状态。例如，堑壕战博弈中，（击中，击中）并非在所有的阶段都是最佳策略，因为有些阶段双方参与者保持休战状态（不击中，不击中）才是最佳策略。接着，考虑到无限次重复博弈及对应的策略所产生的收益让参与者选择了那套普遍意义上的最佳策略。我们知道如此行为并不会形成阶段博弈的纳什均衡，这就意味着一个参与者通过不寻常的行为能够增加他的当前收益。也就是说，违背协议可以获得短期收益。现在，阻止违背协议的行为发生的唯一方法就是用较低的未来收益威胁参与者。为了让威胁生效，未来损失的收益必须足够高，这就需要参与者足够重视未来收益。换句话说，δ 必须足够大。如果 $\delta=0$，那么参与者不会关心未来收益，只是专注于当前收益。此时他们确实会违背协议，所以合作关系是不稳定的。这一推测可导出如下的一般结论。


点拨

不确定重复博弈或者无限次重复博弈，未来总有交手的可能。如果参与者足够关心他们的未来福利并且未来也有足够大的交手的可能性，那么此时就达到了合作性均衡。

为了使合作关系足够稳定，必须满足一些条件。首先，遭遇必须是重复的，并且未来总有遭遇的可能。其次，参与者如何行动必须具有可知性，我们假设这场博弈的历史是共同知识，其实就隐含了这个条件。当且仅当违背协议的行为是可以被知晓的，进而可以被处罚，此时惩罚机制才能得以运行。最后，参与者必须足够关心未来所发生的事情及对自己收益的影响。

13.3 理解测验

无限次重复博弈是以图13—11中的阶段博弈为基础的。设想一个对称的策略收益中，参与者最初选择x，并且只要没有参与者选择y，他将继续选择x；一旦有人选择y，他随后会选择z。求折现系数要满足什么条件，上述策略才成为子博弈精炼纳什均衡。

参与者 2

参与者 1	x	y	z
x	5,5	2,7	1,3
y	7,2	3,3	0,1
z	3,1	1,0	2,2

图 13—11

13.4 针对重复“囚徒困境”的实验证据

当合作关系开始建立时，博弈论提供了一些相当精确的预言。思考图 13—12 所示的“囚徒困境”，这场博弈有唯一的纳什均衡，在均衡中双方参与者各自为战。当然，如果双方参与者开始友好的合作，他们可以将收益从 2 增加到 3，因此这里也有合作的可能。

概述本章的分析结果，博弈论做出如下预言：

预言 1：如果“囚徒困境”只进行一次，参与者会选择各自为战。

预言 2：如果“囚徒困境”进行的次数是有限的，参与者每一阶段都会选择各自为战。

预言 3：如果“囚徒困境”存在的次数是无限次的或者是不确定的，参与者在一些阶段有可能会选择友好合作。

预言 4：如果囚徒困境存在的次数是无限次的或者是不确定的，当这场博弈持续下去的可能性（或者折现系数）很高时，参与者更有可能选择友好合作。

	参与者 2		
		各自为战	友好合作
参与者 1	各自为战	2,2	4,1
	友好合作	1,4	3,3

图 13—12　囚徒困境

以上预言只是源于理论，然而这是否真正符合人类的行为逻辑呢？世界各国的大学生对这一博弈进行过一次实验，实验观察者近距离地观察并记录下了他们的行为。实验结果并没能很好的证明理论，然而，也并不是让人完全失去了信心。这个实验为一些预言提供了事实依据，而其他预言却遭到了有力的反驳。我猜想许多关于大学生的行为我们依然无法理解（包括他们究竟在 Facebook 上上传了什么）。

这项研究中最具代表性的是对加州大学洛杉矶分校的 390 名大学生进行的一系列实验。这些学生被请求参加为期一个回合的囚徒困境或者有限次重复的囚徒困境（为期两个阶段或四个阶段），或者不确定视野的囚徒困境。在每一局博弈的最后，通过是否有参与者出局决定这场博弈是否结束，只要未结束，不确定视野的博弈就一直如此继续下去。在一次实验中，博弈继续下去的概率（我们用 p 表示）为 50%（我们希望这场博弈仅进行两个回合，尽管按道理它可以永远继续下去）；在其他实验中，概率为 75%（我们希望博弈进行 4 个回合）。每个案例中，在所有参与者都保持匿名的状态下，两个学生随机组合成比赛对手，所有博弈的相关信息都是提前知晓的。我们以真实的货币做赌注，假设图 13—12 中的收益是美元。通过一个小时的工作，所有学生的报酬在 12 美元到 22 美元之间浮动，但每个学生平均到手的报酬为 19 美元。

学生们友好合作关系的概率如表 13—3 所示。与预言 1 相反，在为期一个回合的博弈方案中，参与者有 9% 的概率会选择友好合作。但是依然有 91% 的概率参与者的选择与预言保持一致。更多的问题在于预言 2，当学生们知道他们仅有两个回合的对决机会时，他们在第一回合选择友好合作的概率仅为 13%；当他们知道他们会有 4 个回合的对决机会时，他们在第一回合合作的概率会增大到 35%。

表13—3　每个回合参与者合作的概率

	1	2	3	4	5	6	7	8	9	10	11	12
一回合	9%											
T=2	13%	7%										
T=4	35%	22%	19%	11%								
$p=\frac{1}{2}$	31%	26%	20%	13%	13%							
$p=\frac{3}{4}$	46%	41%	39%	35%	33%	27%	25%	26%	29%	26%	32%	31%

尽管这是一个令人失望的事实依据，但是这个实验还是有一些积极因素的。首先，在一场有限次重复博弈的最后阶段，参与者的行为与为期一个回合的博弈中参与者的行为几乎相同。为期一个回合的博弈中，参与者合作的概率为9%；两个回合的博弈中，后一个回合中参与者合作的概率为7%；四个回合的博弈中最后一个回合参与者合作的概率为11%。通过比较这些数字我们得知，学生在有限次重复博弈的最后一个回合的选择与他们在为期一个回合的博弈中的选择相同。而且，随着博弈进入尾声，参与者的合作次数下降。尽管任何存在合作关系的博弈都是一个谜，但是还是有少量的博弈，持续周期越短参与者受处罚的机会就越小。

我们再来讨论不确定视野的情况（根据相关理论，这与无限次视野的情况大致相同）。我们回想起一个问题，博弈的实际视野是不确定的，它会随着实验的不同而有不同的视野。当$p=\frac{1}{2}$时，实验最长持续5个阶段，而当$p=\frac{3}{4}$时，最长持续12个阶段。与预言3相同，博弈中存在着大量合作关系。更有趣的是关于有限次视野与不确定视野的比较结果。理论显示不确定视野中应该有更多的合作，实验结果也确实证实了这一点。我们注意到，从第一阶段的角度来看，$T=2$和$p=\frac{1}{2}$时，博弈的持续时间是相同的，所以参与者也希望在选择做出不友好的行为时得到相同的惩罚。在一场不确定视野的博弈中，参与者友好合作的概率为31%，而在一场有限次博弈中，维持合作关系的概率仅为13%。通过对比$T=4$及$p=\frac{3}{4}$时的概率，我们可以得出类似的结果。最后，与预言4相同，博弈持续下去的可能性越大，参与者之间合作的概率越高。如果参与者持续遭遇的概率为75%，那么学生选择友好合作的概率在25%到46%之间。相反，如果博弈持续下去的概率仅为50%，参与者合作的概率仅维持在13%到31%之间。

尽管在实验中，即使是为期一个回合的博弈，同学们也会相互合作完成任务，但是在现实社会中真的能够如此吗？当有更多的美元作为利益上的诱惑时他们还能如此吗？是的，他们确实是那样做的。有一个真人秀节目叫做《朋友还是敌人》（*Friend or Foe*），其中有一个环节让参与者做类似于“囚徒困境”的一个游戏。两个参与者组合成对手，在最初的时候一起合作，回答一些琐碎的问题。正确回答一个问题将会产生500美元或1 000美元的“信任基金”。第二个阶段，他们开始一场名为“囚徒困境”的博弈，这场博弈所需的大部分资金来源于“信任基金”。

我们用V表示信任基金，用图13—13描述他们面对的策略情形（所填入的就是以货币形式表示的收益）。参与者同时在标记为“朋友”和“敌人”的策略中做出选择。如果双方都选择“朋友”，他们一起分享信任基金；如果一方选择“朋友”另一方选择“敌人”，后者将独享信任基金；如果双方都选择“敌人”，他们都得不到信任基金。与囚徒困境相反，“敌人”弱占优。

参与者2

参与者1		敌人	朋友
	敌人	0,0	V,0
	朋友	0,V	$\frac{V}{2}$,$\frac{V}{2}$

图13—13　朋友或者仇敌

表13—4是源于这项研究的39个片段中的发现。第一行为“总计”，表示数据的加总。信任基金的平均值为3 705美元。尽管有如此高的收益作为诱惑，参与者选择“朋友”的概率仅为50%，如果每个参与者都希望其他所有人选择“朋友”的概率为50%，那么选择合作将产生926.25美元的期望收益，而选择“敌人”得到的收益是前者的两倍，即1 852.5美元。于是，选择这个弱占优策略需要参与者支付超过900美元的成本。

为什么参与者能够忽视金钱成本保持合作关系呢？解决这一问题的线索在于弄清楚参与者的个性如何影响他们的行为。如表13—4所示，女性之间合作的概率高于男性，白种人之间合作的概率高于非白种人，老年人合作的概率高于年轻人。你如何理解这种现象呢？

表13—4　个性特点与“朋友或敌人”策略中的合作博弈

	琐碎收入（美元）	合作概率	实得收入（美元）
总计	3 705	50%	1 455
男性	4 247	45%	1 834
女性	3 183	56%	1 088
白种人	3 957	53%	1 417
非白种人	2 825	42%	1 587
年轻人（<31岁）	3 603	41%	1 592
成年人（≥31岁）	3 839	63%	1 276

但是关于策略的谋划并没有到此结束。一场博弈不单取决于参与者自身的个性特点，也与其合作或竞争对手的个性特点有着密切的关系。与实验相反，在《朋友还是敌人》中一个参与者会揣测其伙伴的意图。表13—5显示了基于双方个性特点的参与者选择合作的概率。得到这些数据的第一步是了解横排参与者选择合作的概率。比如，当两个异性组合成一组时，男性合作的概率为43%而女性合作的概率为55%。

表13—5　合作伙伴的个性特点与"朋友或敌人"中的合作博弈

	男性	女性
男性	48%,48%	43%,55%
女性	55%,43%	56%,56%
	年轻人	成年人
年轻人	40%,40%	42%,63%
成年人	63%,42%	63%,63%
	白种人	非白种人
白种人	51%,51%	58%,44%
非白种人	44%,58%	25%,25%

女性的合作伙伴的性别对她的博弈没有任何影响，因为55%（女性与男性合作）与56%（女性与女性合作）没有多大差别。作为对照，男性与女性合作的概率（43%）低于与男性合作的概率（48%）。合作伙伴的年龄也与此没有多大关系。最显著的要素在于种族：不同种族间的合作概率高于种族内的合作概率。当一个白种人与非白种人组合时，他与之合作的概率为58%，而与另一个白种人合作的概率仅为51%。一个非白种人与白种人合作的概率为44%，而与另一个非白种人合作的概率仅为25%。不仅如此，许多人类行为的奥秘还有待我们去揭开！

谜题

为什么在一个回合的"囚徒困境"中会有参与者以合作的方式开始博弈呢？在有限次重复的囚徒困境中，如果不能保持合作关系，参与者会感到不安吗？如果会，原因何在？参与者的合作会不循惯例吗？因为许多现实中的遭遇与未来遭遇的概率并不一致。在我们的模型中有什么遗漏吗？

本章总结与练习题

本章总结

本章探讨了在共同维持诱人的收益这一意识的指导下，人们实现**合作性均衡**的能力。我们的分析以一个残酷的策略场景开始——第一次世界大战中的堑壕战，并不是对每个战士而言纳什均衡都是最佳选择。如果所有的士兵都放弃杀死敌人的念头，他们最好选择停火。特别值得一提的是，每个士兵都在尽最大努力向敌人开枪（的确，这是一个纳什均衡），但是这样做的结果是每一方都要经历一段悲惨的遭遇，所以他们最好选择休战。挑战在于如何让休战协议永久地发挥作用以使得所有士兵都能遵守。

充实这一策略的决定性要素在于把它设计成一场**重复博弈**以给参与者（堑壕里的士兵）提供未来遭遇的机会。维持合作关系的关键在于将来是否有遭遇的机会，这个机会或者是确定的（**无限次视野**）或者是有某种可能的（**不确定视野**）。但是，即使一场博弈是重复的，惠及所有参与者的结果（比如第一次世界大战期间维持和平的休战协定）也会诱惑参与者背叛协议，因为这样做能增加他当前阶段的收益。这是一个必然的结果，因为这一结果没能形成阶段博弈的纳什均衡。要维持稳固的合作关系，需要一些压力迫使参与者不因短期博弈中利益的诱惑而去背叛合作协议。这种压力就是参与者能够预料到背叛行为会影响到未来行为，降低参与者的未来收益。在这场博弈以后的阶段里，以**"冷酷策略"**为指导，参与者行为上转向阶段博弈的纳什均衡。博弈的重复为设计一套维持合作关系的奖惩机制提供了可能。如果所有的参与者在博弈中友好合作，作为奖励他们在未来阶段会继续这种合作关系并能获得高额收益。但是，如果一个参与者为了增加他的当前收益而损害他人的利益，背叛合作协议，这个背叛者将得到惩罚，获得较低的收

益。在本章的分析中，似乎所有的参与者都会因背叛行为遭到惩罚（因为合作关系破裂），但是在下一章我们就会明白只有那些不善伪装的参与者才会遭受惩罚。

到目前为止，上面的论述集中在维持合作的可能性：存在一个维持合作关系的子博弈精炼纳什均衡。同时，也有的均衡缺少合作关系，比如在这样的博弈中，参与者只是根据阶段博弈的纳什均衡重复同样的行为。如果参与者当前处于均衡中，他们的行为将导致低收益，参与者怎样才能把它转变成一个较好的均衡呢？说得更明确一点，在法国的堑壕里，战场上的士兵怎样做才能不受这一准则的束缚，在休战协定发挥作用期间试图杀死敌人？在这种情况下，如果有人撕毁休战协定，他会受到报复的威胁。这是一个有一定难度但是却很关键的问题，关于怎样做才能切实地达成合作关系这一问题，我们依然没能找到很好的答案。博弈论在这个难题面前也无能为力。

博弈论能解释的只是合作什么时候可行，什么时候不可行。这个让合作关系有效运行的机制只有在特定的条件下才能够发挥作用。首先，参与者在未来的博弈中必须有对手。只有在未来惩罚机制的威胁下，违背合作原则的行为才能被阻止，但是让这一威胁生效的必要条件是保证博弈一直持续下去。其次，参与者必须充分衡量其未来收益，正如在接近 1 的折现系数中所反映的那样：如果一个参与者的目光不够远，他只是高度关心自身的当前收益，那么当他可以通过某种行为增加当前收益时，未来收益降低的惩罚也阻止不了他。因此，参与者必须有足够的耐心，充分衡量其未来收益。想让合作关系稳固，参与者必须有“工作的耐心”，不去信奉贺瑞斯（Horace）的名言：“活在当下，及时行乐，不相信明天。”

练习题

1. 假设有三个渔夫，他们每天各自决定派出去多少艘船到当地的湖里捕鱼。一个渔夫可以派出一艘或两艘船，每艘船平均每天的成本为 15 美元。出海的船越多捕到的鱼越多。但是，每天捕获的鱼的数量是有限的，一个渔夫派出去的船越多，剩下的渔夫捕到的鱼越少。表 13—6 表明渔夫捕到的鱼的数量取决于每个渔夫派出海的捕鱼船的数量。

渔夫的当前收益是他捕获到的鱼的价值（假设每条鱼售价 1），减去渔船的成本。例如，一个渔夫派两艘船出海，另外两个渔夫各派一艘船出海，那么这个渔夫的收益为 75–30=45。因为阶段博弈是对称的，所以表 13—6 是通用的，可以用于计算任何渔夫的收益。渔夫之间进行着一场无限次重复博弈，在每一个博弈阶段，他们同时决定派多少艘船出海。每个渔夫的收益是他所得收益的现值，渔夫 i 的折现系数为 δ_i。寻找一个行动的集合（每一个行动针对一个参与者），使得参与者的收益高于其在阶段博弈的纳什均衡中所得的收益。接着产生了一个“冷酷策略”，为那些行动创造了可能条件。为这个策略寻找条件使其成为一个子博弈精炼纳什均衡。

表13—6　渔夫的捕鱼量及收益

派出的船的数量	其他两个渔夫派出的船的数量	捕到的鱼	船的成本	收益
1	2	40	15	25
1	3	35	15	20
1	4	30	15	15
2	2	75	30	45
2	3	65	30	35
2	4	50	30	20

2. 考虑图 13—14 中关于两个参与者阶段博弈的无限次重复的描述。图中的第一个数字是参与者 1 单个阶段的收益。

参与者 2

参与者 1		v	w	x	y	z
	a	1,1	0,4	0,3	3,2	2,5
	b	4,0	3,3	8,0	1,1	1,2
	c	3,0	0,8	7,7	2,5	0,3
	d	2,3	1,1	5,2	6,6	1,2
	e	5,2	2,1	3,0	2,1	0,0

图 13—14

假设过去的行为是共同知识。每个参与者的收益是单个阶段所得收益的现值，折现系数为 δ。

a. 找到使下列策略成为一个子博弈精炼纳什均衡的条件。

■ 参与者 1：阶段 1，选择 c。阶段 t（$\geqslant 2$），如果第 t-1 阶段的收益为（c，x）选择 c，否则选择 d。

■ 参与者 2：阶段 1，选择 x。阶段 t（$\geqslant 2$），如果第 t-1 阶段的收益为（c，x）选择 x，否则选择 y。

b. 找到使下列策略成为一个子博弈精炼纳什均衡的条件。

■ 参与者 1：阶段 1，选择 c。阶段 t（$\geqslant 2$），（1）如果从第一个阶段到 t-1 阶段的收益为（c，x）选择 c；（2）如果从第一个阶段到 t-2 阶段的收益为（c，x），并且在 t-1 阶段参与者 1 选择 c，参与者 2 不选择 x，那么选择 e；（3）如果从第一个阶段到 t-2 阶段的收益为（c,x）在 t-1 阶段，参与者 1 不选择 c，参与者 2 不选择 x，那么选择 a；（4）以上情况都不成立时选择 b。

■ 参与者 2：阶段 1，选择 x。阶段 t（$\geqslant 2$），（1）如果从第一个阶段到 t-1 阶段的收益为（c，x）选择 x；（2）如果从第一个阶段到 t-2 阶段的收益为（c，x），且在 t-1 阶段，参与者 1 选择 c，参与者 2 不选择 x，那么选择 v；（3）如果从第一个阶段到 t-2 阶段的收益为（c，x），且在 t-1 阶段，参与者 1 不选择 c，参与者 2 选择 x，那么选择 z；（4）以上情况都不成立时选择 w。

3. 假设图 13—14 中的阶段博弈重复了三次并且每个参与者的收益是他们每个单一阶段的收益之和。假设参与者过去的行为是共同知识。判断下列策略是否是子博弈精炼纳什均衡。

■ 参与者 1：阶段 1，选择 c。阶段 2，如果阶段 1 的收益为（c，x）选择 c，否则选择 b。阶段 3，如果阶段 1 和阶段 2 的收益为（c，x）选择 d，否则选择 b。

■ 参与者 2：阶段 1，选择 x。阶段 2，如果阶段 1 的收益为（c，x）选择 x，否则选择 w。阶段 3，如果阶段 1 和阶段 2 的收益为（c，x）选择 y，否则选择 w。

4. 考虑图 13—15 中关于两个参与者阶段博弈的无限次重复的描述。表中的第一个数字是参与者 1 单个阶段的收益。

参与者 2

参与者 1		w	x	y	z
	a	2,7	1,4	1,3	5,6
	b	1,1	2,2	0,4	3,3
	c	3,2	4,3	3,1	8,0
	d	4,0	3,1	5,1	2,3

图 13—15

假设参与者过去的行为是共同知识。每个参与者的收益是单一阶段所得收益的现值，参与者 i 的折现系数为 δ_i。

a. 定义一个"冷酷策略"。

b. 找到使 a 中的策略是子博弈精炼纳什均衡的条件。

5. 考虑关于图 13—16 中阶段博弈的无限次重复的描述。假设每个参与者的收益是他所得利润的现值，折现系数为 δ。

参与者 2

参与者 1		w	x	y	z
	w	4,4	5,3	6,2	1,1
	x	3,5	6,6	7,10	2,7
	y	2,6	10,7	8,8	2,11
	z	1,1	7,2	11,2	3,3

图 13—16

a. 寻找一个策略以便产生如下结果，每个阶段两个参与者都选择 x，并且这个策略是子博弈精炼纳什均衡。

b. 寻找一个策略以便产生这样的结果，两个参与者在每个单数阶段都选择 x，偶数阶段选择 y，这个策略是子博弈精炼纳什均衡。

c. 寻找一个策略以便产生这样的结果，两个参与者从阶段 1 到阶段 10 选择 x，之后选择 z，这个策略是子

博弈精炼纳什均衡。

d. 假设 $\delta=\frac{2}{5}$，寻找一个策略以便产生以下结果，两个参与者每个阶段都选择 y，并且这个策略是子博弈精炼纳什均衡。

6. 维持合作关系的一个众所周知的策略就是“以牙还牙”。参与者在博弈开始时保持合作关系，之后根据对手前期的选择做出相应的行为。“以牙还牙”包含着“以其人之道还治其人之身”的思想。一场堑壕战博弈以下列形式进行：阶段 1 选择不击中敌人。阶段 t（≥ 2），如果另一个参与者前期选择不击中对方则不击中，如果另一个参与者前期选择杀死敌人则杀死对方。对于一场无限次重复的堑壕战博弈，为“以牙还牙”寻找条件使其成为子博弈精炼纳什均衡。

7. 我们将参与者的收益设定为 X，参与者 1 的策略是为了得到价值为 x_1 的收益。类似的，参与者 2 的策略是为了得到价值为 x_2 的收益。参与者 1 的收益为 v_1（x_1，x_2）$=5+x_1-2x_2$，参与者 2 的收益为 v_2（x_1，x_2）$=5+x_2-2x_1$。

a. 假设 X 的取值范围在 1 到 4 之间（包括 1 和 4）。（值得注意的是，取值不仅限于整数，还包括诸如 2.648 和 1.00 037 这样的小数）。求出所有的纳什均衡。

b. 现在我们假设这场博弈是无限次持续下去的，参与者的收益是他单个阶段所得收益的现值，折现系数为 δ。

（1）假设 X 仅包括两个收益：2 和 3。因此，参与者只能在 2 和 3 之间做出选择而不能选择其他的收益。考虑以下策略：阶段 1，参与者选择收益 2。阶段 t（≥ 2），如果两个参与者前期都选择 2，那么这个参与者选择 2 作为收益；否则他选择 3。求出保证这是一个子博弈精炼纳什均衡的条件。

（2）我们现在假设 X 在 1 到 4 之间取值，所以参与者可以在 1 到 4 之间（包括 1 和 4）任意选择。考虑以下策略：阶段 1，参与者选择 y。阶段 t（≥ 2），如果两个参与者前期都选择 y，那么这个参与者选择 y，否则选择 z。y 和 z 都是源于 X 中的数据。更进一步，假设 $1 \leq y \leq z \leq 4$。找到关于 y，z 和 δ 满足的条件，使其成为一个子博弈精炼纳什均衡。

8. 有三个以上的博士组成一个团队，每个阶段，这些博士都要决定以何种态度努力工作。将博士 i 在阶段 t 做出的选择表示成 e_i^t，假设 e_i^t 可以在 1 到 10 这 10 个等级中选择。这些参与者工作越努力，整个团队得到的收益越高。更明确的，整个团队的总收益等于总体努力程度的两倍：收益 $=2\times$（$e_1^t+e_2^t+\cdots+e_n^t$）。

博士的收益等于他的总收益减去个人付出的努力。因此：博士的收益 $=(\frac{1}{n})\times 2\times(e_1^t+e_2^t+\cdots+e_n^t)-e_i^t$

这个阶段博弈是无限次重复的，每个博士的收益就是他所得收益的现值，且博士 i 的折现系数为 δ_i。

a. 假设博弈的前期结果是共同知识。也就是说，在阶段 t，所有的博士从阶段 1 到阶段 $t-1$ 这些过去所做的选择都是可观察的。求出一个子博弈精炼纳什均衡，每个参与者选择的努力程度 $e^* > 1$。

b. 假设博弈的前期结果是不透明的。在阶段 t，只有参与者的总努力（$e_1^\tau+e_2^\tau+\cdots+e_n^\tau$）在阶段是被所有参与者共同知晓的（$\tau \leq t-1$）。（我们假设参与者的回忆是完美的，参与者知道他自己过去的努力，不过你可以忽略这个信息。）寻找一个子博弈精炼纳什均衡，每个参与者选择的努力 $e^* > 1$。

9. 下面引用的内容源于丘吉尔首相的备忘录，而这些内容是他在第二次世界大战的最后一年写给一个将军的。

> 我希望你严肃认真地考虑毒气战这一问题。在第一次世界大战中，每个士兵都使用毒气作战，这一行为居然没有招致来自道义上或者宗教上的谴责，从道义方面讨论这个话题荒谬至极。另外，第一次世界大战中对城市的轰炸是被禁止的。现在人们却毫无顾忌地这样做。这仅仅是一个流行趋势改变的问题，就像女士们根据自己的喜好改变裙子的长短。为什么德军不曾使用过毒气呢？我们很难确定这种行为是否出于其良心上的顾忌或者偏好。他们使用毒气战是因为这对他们来说无利可图……他们不对我们使用毒气战的唯一原因在于他们害怕遭到报复。能对他们造成伤害的就是对我们有利的。我希望明智的人们以一种理智的态度研究这个问题。

我们考虑一下丘吉尔和希特勒面临的情形。假设每一阶段他们都要决定用多少毒气作战。我们把丘吉尔在阶段 t 使用的毒气量用 g_c^t 表示，用 g_h^t 表示希特勒在阶段 t 使

PDG

用的毒气量。假设每一阶段，丘吉尔可以在0和G_c之间选择要使用的毒气量：$0 \leqslant g_c^t \leqslant G_c$。希特勒可以在0和$G_h$之间选择：$0 \leqslant g_h^t \leqslant G_h$。阶段$t$，丘吉尔所得收益为

$$V_c(g_c^t, g_h^t) = ag_c^t - bg_h^t$$

希特勒所得收益为

$$V_h(g_c^t, g_h^t) = dg_h^t - eg_c^t$$

我们假设$b > a > 0$，$e > d > 0$，$G_c > 0$，$G_h > 0$。所以阶段t，丘吉尔所得收益会随着他在军事上所用毒气量的增加而增加（可能是因为这样增加了胜算的可能），随着德方使用毒气量的增加而减少（原因很明显）。我们用类似的方法讨论希特勒的收益。假设每一阶段战争结束的概率为p（$0 < p < 1$）。如果博弈结束，每个参与者的收益都是零。那么丘吉尔的收益可以表示成

$$V_c(g_c^1, g_h^1) + pV_c(g_c^2, g_h^2) + p^2V_c(g_c^3, g_h^3) + \cdots$$

或$\sum_{t=1}^{\infty} p^{t-1} V_c(g_c^t, g_h^t)$

类似的，希特勒的收益表示成：

$$V_h = (g_c^1 + g_h^1) + pV_h(g_c^2, g_h^2) + p^2V_h(g_c^3, g_h^3) + \cdots$$

或$\sum_{t=1}^{\infty} p^{t-1} V_h(g_c^t, g_h^t)$

每一阶段，两个参与者同时开始行动，他们知道整个过去的选择。回答以下问题需要理智的头脑：

a. 设计一个策略及限制因素使其成为一个子博弈精炼纳什均衡，并且确保没有国家使用毒气作战；这个均衡结果需要满足$g_c^t = 0$，$g_h^t = 0$，$t = 1,2,3,\cdots\cdots$

b. 设计一个策略及限制因素使其成为一个子博弈精炼纳什均衡，并且确保两个国家都使用毒气作战；这个均衡结果需要满足$g_c^t > 0$，$g_h^t > 0$，$t = 1,2,3,\cdots\cdots$

c. 设计一个策略及限制因素使其成为一个子博弈精炼纳什均衡，并且确保丘吉尔使用毒气但是希特勒不使用毒气；这个均衡结果需要满足$g_c^t > 0$，$g_h^t > 0$，$t = 1,2,3,\cdots\cdots$

附录 1

所得收益的现值

假设你很富有，你获悉一个脾气古怪的伯父去世了，令你意想不到的是，他在遗嘱中提到将部分遗产留给你。由于他的古怪性格，他留给你一个选择。在遗嘱中列着三组五年内将得到的收益，用选项 A,B,C 表示（如表 13—7 所示）。根据老人的遗愿，你只能选择其中的一个计划。

表13—7　收益

年份	选项A	选项B	选项C
1	100	40	30
2	80	50	40
3	70	60	50
4	40	100	80
5	30	120	100

考虑到你伯父是一个有趣而古怪的人，你开始分析遗嘱中的三个选项。你很快决定删除选项 C：因为选项 B 在五年内的每一年都能得到很多钱，每个爱财的人都倾向于选 B 而不是 C。然后你会疑惑地问自己："该死，难道我伯父认为我是个笨蛋吗？"你决定不选择 C。不幸的是，选项 A 和选项 B 之间的比较并不简单：选项 A 在阶段 1,2,3，能够得到比较多的钱，但是选项 B 在阶段 4,5 能够得到较多的钱。我们现在需要找到一个方法比较选项 A 和 B 的收益。

比较金钱（效用或者收益）的一个通用方法就是我们所熟知的现值。我们统一用货币流去衡量，方法就是努力获得一笔钱，如果今天你得到这笔钱（之后会两手空空），那么你会对这笔钱及其价值漠不关心。换句话说，这笔钱的现值何在?

为了明确这笔钱的现值，你需要问下自己："第二年 100 美元的收益换成今天的价值能有多少？" 也就是，如果你将在得到 Y 的收益及第二年 100 美元的收益之间做出选择，Y 的何种价值会让你保持中立呢？为了回答这个问题，假设你在今后的一年内不打算花这笔钱。在这种情况下，如果你现在得到的收益为 Y，你会把它存入银行得到利率为 r 的利息。如果利率为 15%，那么 r=0.15。因此，如果现在你得到的收益为 Y，明天你将得到的收益为 $1.15\times Y$。为使收益相同，必须保证下列式子成立：

$$1.15\times Y=100\Rightarrow Y=\frac{100}{1.15}=87$$

所以，今天得到 87 美元的收益与第二年得到 100 美元的收益是相等的。

现在假设你到第三年得到 100 美元的收益。如果现在你得到的收益为 Z，你将这笔钱存入银行两年。经过一年，你得到的收益为 $(1+r)\times Z$，你把这笔钱继续存入银行，两年后的年末得到的收益为 $(1+r)\times(1+r)\times Z$，或者表示成 $(1+r)^2\times Z$。这笔钱的现值为：

$$1.15^2\times Z=100\Rightarrow 1.322\,5\times Z=100\Rightarrow Z=\frac{100}{1.322\,5}=76$$

所以，现在得到 76 美元的收益与两年后得到 100 美元的收益是相同的。

说得更普遍一点，现在得到 W 的收益与 t 年后得到 1 美元的收益是相同的，W 满足以下等式：

$$(1+r)^t\times W=1\Rightarrow W=\frac{1}{(1+r)^t}=\left(\frac{1}{1+r}\right)^t$$

因此，t 阶段的 1 美元与今天的 $[1/(1+r)]^t$ 美元是等价的。在这个方程式中，r 表示折扣率，（为了简化表达）用 δ 表示 $1/(1+r)$，δ 为折现系数。

为了计算这笔钱的现值，我们把所有的未来收益转变成收益现值。我们刚刚证明了某一阶段的 1 美元与现在的 δ 美元是等价的，所以明天的收益需要用 δ 衡量。类似的，两个阶段得到的收益需要用 δ^2 衡量，以此类推。假设从阶段 1 到阶段 T 的收益分别为 $u_1, u_2, u_3, \cdots, u_T$，它们的现值表示为：

$$u_1+\delta u_2+\delta^2 u_3+\delta^3 u_4+\cdots+\delta^{T-1}u_T$$

当前收益就是带权重单个收益之和。我们注意到这个衡量单位是递减的，由于 $0<\delta<1$，每次 δ 乘以它自身，值就会减小。这说明随着时间的推移，收益的价值在逐渐降低，人类的行为普遍表现出了这一特性：人类缺乏耐心！

让我们用这个方法进行一下测试，即当折扣率为 r=0.10（同时也包含着 δ=0.91）时，计算选项 A 和 B 的现值：

选项 A：$100+(0.91\times 80)+(0.91^2\times 70)+(0.91^3\times 40)+(0.91^4\times 30)$

$=100+(0.91\times 80)+(0.828\times 70)+(0.754\times 40)+(0.686\times 30)=281.50$

选项 B：$40+(0.91\times 50)+(0.91^2\times 60)+(0.91^3\times 100)+(0.91^4\times 120)=292.83$

选项 B 的现值大于选项 A 的现值，因此，根据这一比较结果，你更倾向于选择 B 而不是 A。但是，如果你不更多地考虑未来收益，例如，令 r=0.25, 那么 δ=0.80，选项 A 优于选项 B：

选项 A：$100+(0.8\times 80)+(0.8^2\times 70)+(0.8^3\times 40)+(0.8^4\times 30)$

$=100+(0.8\times 80)+(0.64\times 70)+(0.512\times 40)+(0.409\,6\times 30)=241.57$

选项 B：$40+(0.8\times 50)+(0.8^2\times 60)+(0.8^3\times 100)+(0.8^4\times 120)=218.75$

据前面的分析，选项 A 创造的收益更高，当你给未来收益更小的权重，即你更没有耐心时，选项 A 更具吸引力。

如果收益是无限的结果又会怎样呢？那么现值为：

$$V=u_1+\delta u_2+\delta^2 u_3+\delta^3 u_4+\cdots=\sum_{t=1}^{\infty}\delta^{t-1}u_t \qquad (13.6)$$

尽管这样无限阶段数字的加总是一个令人厌倦的计算，但是当收益固定时，计算就相当简单了。如果 $u_i=u$，所有阶段相加：

$$V=u+\delta u+\delta^2u+\delta^3u+\cdots \quad (13.7)$$

为了简化 V 的表达式，将等式（13.7）的两边都乘以 δ：

$$\delta V=\delta u+\delta^2u+\delta^3u+\delta^4u+\cdots \quad (13.8)$$

现在用式（13.8）减去式（13.7）：

$$V-\delta V=(u+\delta u+\delta^2u+\delta^3u+\cdots)-(\delta u+\delta^2u+\delta^3u+\delta^4u+\cdots)$$

右边的式子可以通过简化删去一些表达式；两个表达式中都包含 δu，δ^2u，δ^3u 可以删除这些相同的式子简化最终表达。实际上，除了 u，右边所有的式子都能被删除。因此，我们得到下面表达式：

$$V-\delta V=u\Rightarrow V(1-\delta)=u\Rightarrow V=\frac{u}{1-\delta}$$

是不是很神奇，我们把无限表达转化成了有限表达，最终简化成一个字母。

现在我们假设你那位很有幽默感的伯父给了你第四种选择，用 D 表示，据这一选项，每年能得到 30 美元（假设你会用无限长的寿命享受这笔财富）。如果 $\delta=0.91$，据我们之前的分析，选项 D 的现值为 30/(1−0.91）=333.33，这个结果要优于其他三个选项。

假设你得到的收益是无限次重复且恒久不变的，但是这一假设是从 T 阶段开始的。把这一收益折算成现值，我们发现现值为：

$$0+\delta\times 0+\cdots+\delta^{T-1}\times 0+\delta^{T}u+\delta^{T+1}u+\delta^{T+2}u+\delta^{T+3}u+\cdots=\delta^{T}(u+\delta u+\delta^2u+\cdots)=\delta^{T}\left(\frac{u}{1-\delta}\right)$$

一旦到了阶段T，你在阶段T所得收益的价值为$u/(1-\delta)$，但是把它折算成阶段1的收益，现值仅为$\delta^{T}[u/(1-\delta)]$。

这个计算技巧允许我们把无限个数字的加总转化成一个数字，只要所得收益是有周期性的，这一转化就能进行因为有周期性的收益是自身的不断重复（我们注意到每个阶段恒定不变的数字都是自身的重复。）比如，奇数阶段用 x 表示，偶数阶段用 y 表示：

$$V=x+\delta y+\delta^2x+\delta^3y+\delta^4x+\delta^5y+\cdots$$

把上述式子合并组合，得到以下结果：

$$V=x+\delta y+\delta^2(x+\delta y)+\delta^4(x+\delta y)+\cdots \quad (13.9)$$

等式（13.9）可以被认为是每一个其他阶段参与者得到 $x+\delta y$ 的收益。等式两边同时乘以 δ^2（是 δ^2 而不是 δ，因为我们已经对原等式进行了重新排列以使得每一个其他阶段都有所收益）：

$$\delta^2V=\delta^2(x+\delta y)+\delta^4(x+\delta y)+\delta^6(x+\delta y)+\cdots \quad (13.10)$$

现在我们用等式（13.9）减去等式（13.10）：

$$V-\delta^2V=[x+\delta y+\delta^2(x+\delta y)+\delta^4(x+\delta y)+\cdots]-[\delta^2(x+\delta y)+\delta^4(x+\delta y)+\delta^6(x+\delta y)+\cdots]$$

$$V(1-\delta^2)=x+\delta y\Rightarrow V=\frac{x+\delta y}{1-\delta^2}$$

比如，每一个其他年度得到的收益为 25 和 35（从 25 开始），这样的现值为：

$$\frac{25+0.91\times 35}{1-0.91^2}=330.72$$

13.4 理解测验

假设有一串无限的收益组合，阶段1收益为100，阶段2收益为50，阶段3收益为25，接着就是这三个数字的重复：

$$100+\delta\times 50+\delta^2\times 25+\delta^3\times 100+\delta^4\times 50+\delta^5\times 25+\cdots$$

将上述求和表达式转化成一个简单的表达式。

我们最后的任务是去分析一组随机的收益。一场重复博弈也就意味着所得收益是无限的。我们之间存在这样的协议，如果到阶段 t 你依然在参与博弈，那么你得到的收益为 u_t。但是，每一阶段，你能坚持到下一阶段的概率为 p；因此，每一阶段结束时，你牺牲的概率为 $1-p$。让我们计算一下期望的现值，我们认定所谓的价值是由你参与博弈时得到的收益决定的。为了让推导更清晰，我们用 d 而不是 δ 表示折现系数。

这一连串收益的期望现值为：

$$V=u_1+pdu_2+p^2d^2u_3+p^3d^3u_4+\cdots=\sum_{t=1}^{\infty}p^{t-1}d^{t-1}u^t=\sum_{t=1}^{\infty}(pd)^{t-1}u_t \tag{13.11}$$

阶段 1 你一定能得到 u_1 收益。你能坚持到第二阶段的概率为 p，得到的收益为 u_2。经过折算，阶段 2 的期望价值为 $p\times d\times u_2$。（当然，一旦你牺牲你得到的收益为零。）你从第二阶段坚持到第三阶段的概率依然为 p，从第一阶段开始算起，你能够参与第三阶段博弈的概率为 $p\times p$，或者表示成 p^2，这是从阶段 1 到阶段 2 及从阶段 2 到阶段 3 能够参与下一阶段概率的综合。因此，阶段 3 的期望收益为 $p^2\times d^2\times u_3$，以此类推。

无限次视野与不确定视野实际上是一样的，为了验证这一想法，用 δ 代替式（13.11）中的 pd：

$$V=\sum_{t=1}^{\infty}(pd)^{t-1}u_t=\sum_{t=1}^{\infty}\delta^{t-1}u_t$$

这一等式与等式（13.6）是一样的。折现系数 δ 可以由时间偏好（一个参与者会对未来打多少折扣）和博弈结束的可能性共同决定。

附录 2

动态规划

如果作为子博弈精炼纳什均衡的一部分的这个策略是最优的，那么假设其他参与者的策略是既定的，这个策略能够不受信息设定的影响使参与者获得最大收益。在一场重复博弈中，参与者之间知道彼此过去的策略及选择，“每一个信息”也就是“每一阶段参与者的选择”。

定义 13.2

最优策略（Optimal Strategy）：如果每一阶段及过去这个策略都能使参与者获得最大收益，那么这个策略是最优的（作为一个子博弈精炼纳什均衡）。

让我们首先从有可能成为子博弈精炼纳什均衡的一部分的一些策略开始。利用参与者的备选策略计算他的收益是一项非常简单的任务（尽管或许有些乏味）。为了考虑如何完成这一任务，假设现在是阶段 t，我们用 h^t 表示从第一阶段到 $t-1$ 阶段参与者选择的一系列行动。h^t 就是阶段 t 的历史。每一个参与者的备选策略都规定了当前阶段应该采取的行动，可能是 h^t。那一系列行动（一个行动针对一个参与者），我们称之为 a^t，阶段 t 得到的收益可以通过把这一系列行动插到参与者单一阶段的收益中去计算。$t+1$ 阶段的历史是 h^t（阶段 t 之前发生的博弈）及 a^t（阶段 t 发生的博弈）的一个组成部分。给定 h^{t+1}，每个参与者的备选策略规定了参与者 $t+1$ 阶段的一个行为。通过这一系列的行动，我们能够计算 $t+1$ 阶段的收益。重复这一过程，我们能够计算任何备选策略的全部收益（从任何时期，任何历史阶段开始）。

一旦我们得到了一个备选策略的收益，决定这一策略是否为最优策略需要与其他任何可选择的策略所得的收益进行比较（继续假设其他参与者使用他们的备选策略）。但是，这个任务听起来不仅乏味还很难去执行（因为有很多可供选择的策略）。请记住一个策略规定了一个时期，一个历史阶段的一个行为，所以，如果一场博弈是无限次

重复博弈，那么一个策略就是这无限次行动的一个组成部分。

为帮助我们走出困境，早在半世纪以前，数学家理查德·贝尔曼（Richard Bellman）做出了一项有趣的研究发现，称为动态规划。这个方法的第一步是确定我们确实不用对备选策略的收益与每一个可选策略的收益进行比较，而只需要与创造最高收益的策略进行比较。如果备选策略的收益大于等于所有可选策略的最高收益，那么这个策略优于任何其他可供选择的策略。基于这一想法，我们来计算当前阶段每一个可能的行动创造的收益（现在及未来单一阶段收益的现值），同时假设未来阶段参与者的行动是最佳的。确定了未来收益是最高的，我们来计算当前阶段任何行动所产生的最高收益。

这些设想很好，但是我们怎样发现最高的未来收益呢？难道只是努力计算吗？当然不是，这些工作的第一步是证明这个备选策略只是部分最优的。

定义 13.3

部分最优策略（Partially Optimal Strategy）：如果在任何时期，任何历史阶段，与所选择的任何当前行动相比，这个策略规定的行动产生了最高收益，那么这个策略就是部分最优策略（作为一个子博弈精炼纳什均衡）。

如果与任何当前阶段的策略相比较，这个策略是最好的，那么这一策略就是部分最优策略。所以，不管现在的选择如何，将来参与者会按照这个备选策略行事。（请记住未来的一连串行动会随着今天所选择的行动而改变，因为今天一个不同的行动对明天来说就是一个不同的历史，并且一个策略规定了历史上的一个可能行动。）我们已经证明了这个备选策略要优于这一系列有限的备选策略，这也就是为什么我们称之为部分最优策略。

假设我们已经证明了这个备选策略就是部分最优的（并不是很困难，在无限次堑壕战博弈中我们已经做了证明）。假设一段特定的历史，现在参与者处于阶段 t。今天选择其他策略或者未来按照备选策略行事，与上述两种选择相比，这个备选策略创造了最高收益；换句话说，它符合部分最优策略的条件。但是在 $t+1$ 阶段按照备选策略行事确实是最好的选择吗？是的，根据部分最优策略的定义，在 $t+1$ 阶段选择其他备选策略或者未来按照这个备选策略行事，与这两种选择相比，这个备选策略确实是最好的。但是在 $t+2$ 阶段按照这个策略行事依然是最佳选择吗？根据部分最优策略的定义，在 $t+2$ 阶段选择其他备选策略或者未来按照这个备选策略行事，与这两种选择相比，这个备选策略确实是最好的。那么在 $t+n$ 阶段它依然是最优策略吗？我想你已经领会了要点。这一讨论可以应用于未来的任何时期任何历史阶段，因为根据部分最优策略的定义，我们已经证明了在每一时期，每一历史阶段，与选择其他任何行动或者未来按照这个备选策略行事相比，这个备选策略创造了最高收益。


点拨

如果这个策略是部分最优策略，那么他就是最优策略。

为了证明一个策略是最优的，我们只需要证明参与者一旦违背这一策略，哪怕是未来依然按照这一策略行事，他也不可能获得更高的收益。只要我们能够证明参与者不愿违背策略，并且证明不管任何时期任何历史阶段这一选择都是正确的，我们就能得出更加强有力的结论，对任何背叛行为而言，这一策略都是最优的。换句话说，这一策略产生的收益大于等于其他任何策略创造的收益。

第 14 章

合作和名誉：非固定居所参与者重复互动博弈中的应用

学习目标

14.1 惩罚
14.2 礼尚往来
14.3 名誉

本章导读

如果在第一次世界大战的堑壕里，士兵之间能产生合作意识并能长期保持合作关系，那么在其他场合应该也存在着合作。本章列举了一个真实的案例，即野生动物之间存在着合作意识。通过这个案例我们明白合作有多种不同的表现形式，而奖惩机制是合作得以维系的保障。前一章我们提到，为了使合作关系得以长久维持，参与者必须做到不屈服于高额利润的诱惑，不为了获取高额收益进行欺诈行为。为此，我们用未来奖惩机制限制诱惑影响参与者的行为。在奖惩机制的作用下，如果参与者不存在欺骗行为，作为回报他的搭档以后会继续与其合作；相反，一旦他说谎，将来会受到惩罚。

众所周知，竞争会导致低价格，损害所有公司的利益。因此，当这些公司串通一气以较高的价格出售商品的时候，一种实际意义上的合作关系便产生了——尽管这种合作是非法的。我们首先来看一个案例，赫赫有名的克里斯蒂拍卖公司（Christie's）及苏富比拍卖公司（Sotheby's）通过协商达成统一，通过收取客户的费用以获取高额利润。在第 13 章我们已经阐明了这样一种观点，即“使合作得以强制执行的惩罚会导致未来所有合作的终止”，本章我们将讨论几种其他形式的惩罚。现在，我们一起讨论一下上述两家拍卖公司的定价案例，首先假设这两家拍卖公司暂时回到了竞争环境，此时，他们收取较低的费用，获取较低的利润；其次，我们假设一个短暂却激烈的竞争环境，公司收取的费用低于正常竞争环境下收取的费用；最后，我们制定了一套较合理公正的惩罚机制，惩罚违背合作原则的公司。

大多时候，“合作”会惠及所有的参与者。比如，第一次世界大战期间，“合作”产生了暂时的停战，从而惠及了所有的参战者。然而，有些合作并非如此。比如，一个参与者帮助另一个需要帮助的参与者脱离困境，这一现象也称为合作。因为这种援助得到的收益不是立竿见影的，参与者只有预见到他的这种援助未来能得到回报的时候，帮助行为才会产生。**这种相**

互式的合作，也就是"现在你帮助我，将来我会报答你"作为均衡机制的一部分得以长期维持。在 14.2 节我们讲述了这种形式的合作在美国国会及吸血蝙蝠群体中是怎样发挥作用的。

通过名誉这面镜子我们也能明白合作标准的稳定性。一个人与他人合作时，如果他能友好对待其合作伙伴，他将得到一个好名声并且其他合作者也会同样友好对待他；然而一旦他不能友好对待其他合作者，他将被认为是不可靠的，那么他也将得到同样的待遇。在 14.3 节我们通过两个案例探究名誉在维持合作机制时发挥的作用：近代，为什么当地的银行家会贷款给国王；为什么 20 世纪初福特公司支付给员工的工资是当时工资水平的两倍。

合作得以维持的一个重要因素在于任何背离合作原则的行为必会受到惩罚。要使惩罚机制得以运行，当背叛行为发生时其他参与者必须能觉察到，因此他们可以及时惩罚背叛者。尽管我们认为迄今为止，历史上的博弈都是被大家所熟知的，因此，我们可以得知所有参与者过去的行为及他是否曾有过背叛合作原则的记录。不过，并非所有的历史事件都是真实的。14.4 小节提出了一个设想方案，即国家试图控制武装军队。鉴于远距离监视一个国家存在的难度，违背协议的行为不会被发现也不会受到惩罚。我们将重点探讨这样一个协议，即美国和苏联为了限制变动导弹防御系统而制定的协议。这一协议被证明是最有意义的协议并在冷战期间得以完善。

14.1 惩罚

如果我们赦免受罚的人，那么将来会有一些人不再惧怕惩罚。这种纵容让罪恶成长繁殖，同时也成为世上道德缺失现象的标志。因此，宽容是一种脆弱的美德，基督徒更容易具备这一美德，因为基督徒往往容易纵容罪恶而不是抑制罪恶的传播。

——罗伯特•迈克菲•布朗

惩罚对不守规则的参与者具有震慑力，然而，合作得以维系的关键不仅在于惩罚的震慑力，同时也取决于其是否可行。如果惩罚的现实意义在于阻止社会上的不当行为，那么参与者必须相信每个人都能通过努力尽量避免遭到惩罚，即此惩罚机制具有较高的可行性。回顾第 8 章，纳什均衡是以绝对的震慑力为基础的，而子博弈精炼纳什均衡的主要优点在于去除这一震慑力。第 13 章中我们提到，如果德国人选择击中，那么对协约国来说，以同样的方式反击德军将是最佳选择。协约国的最佳策略是贯彻执行这一惩罚措施，因为他们希望德军以同样的方式做出回应——击中，并且协约国一方最好也选择击中。惩罚是一种自我约束，鉴于其他参与者希望认真贯彻这一惩罚措施，每个参与者的最佳策略都是努力避免遭到惩罚。

第一次世界大战期间，惩罚维持了一个和平的停战期，此时的惩罚更为严厉。即使是仅有一次的背叛也会降低士兵的收益，而这一现象与阶段博弈的纳什均衡紧密相关。实际上，合作并非总是不堪一击。本章我们认为，比较严厉的惩罚在维系合作时发挥的作用是短暂的。也许你厌倦了把第一次世界大战作为案例分析，那么让我们从一个新的案例开始，即惩罚在重复博弈中的应用：市场上的价格战争。

14.1.1 价格垄断

竞争对手是我们的朋友，而顾客是我们的敌人。

——ADM 公司的座右铭

国际上有两家最成功的拍卖公司以拍卖艺术品著称：克里斯蒂拍卖公司及苏富比拍卖公司。



这两家公司成立于 18 世纪中期的伦敦，销售过许

多著名的艺术品、古董、美酒、家具和其他奢侈品。这两家拍卖公司最简单的挣钱方法是按销售额向推销员收取费用。如果莫奈的画你卖了 12 000 000 美元，那么售价的 5% 作为所要缴纳的费用，推销员要支付给拍卖公司 600 000 美元。不仅如此，拍卖公司也收取买家的费用。

20 世纪 80 年代后期，艺术品市场异常繁荣，这两家拍卖公司从艺术品市场上获取了高额利润。然而当艺术品市场到达顶峰时，他们的境况严重恶化，生意陷入低谷。1989 年，苏富比拍卖公司获得了 10 亿美元的额外收益，而到了 1992 年，公司却勉强得以维持经营。由于两家拍卖公司间的竞争激化导致业务量减少，收取的费用降低，从而降低了彼此的收益。

为了摆脱困境，克里斯蒂拍卖公司的总裁安东尼•坦南特（Anthony Tennant）与苏富比拍卖公司的总裁阿尔弗雷德•陶布曼（Alfred Taubman）决定进行一场改革，因此他们于 1993 年春在陶布曼总裁伦敦的公寓见面。阿尔弗雷德对安东尼说："如果一直处于这种低迷状态，我们将无法求得生存，我觉得是时候抬高价格了。"这一保守的表态导致这两家拍卖公司在长达七年的时间里相互勾结，人为地抬高收费标准。在这种非法的价格垄断行为被美国当局发现后，苏富比公司支付了 4 500 万美元的罚款，陶布曼总裁在监狱服刑一年，而克里斯蒂公司因其与美国当局的合作关系而免于惩罚。

这两家拍卖公司串通一气进行价格垄断的行为持续了很长时间，对此，用一点儿博弈论的知识就可以解释。为简单起见，我们假设每家公司可以在以下三个数字中选一个作为本公司的收费标准：4%、6%、8%。图 14—1 中描述了这一阶段博弈，它有一个唯一的纳什均衡，在这一均衡的作用下两家公司都会选择 6%。

		苏富比公司		
		4%	6%	8%
克里斯蒂公司	4%	0,0	2,1	4,−1
	6%	1,2	4,4	7,1
	8%	−1,4	1,7	5,5

图 14—1　艺术品拍卖公司的博弈

我们现在假设这场博弈是无限次重复的，在这场博弈中，他们用严峻的触发战略支持高定价：

- 阶段 1，定为 8%；
- 其他阶段：
  如果两家拍卖公司之前定为 8% 则定为 8%；
  如果两家拍卖公司至少有一家之前收取的费用率不是 8%，则定为 6%。

请注意，是否收取 6% 的费用仅取决于前面阶段的收费标准。这一条件有效地维持了合作关系，就像依据过去的博弈制定的惩罚措施更有效一样（正如我们为了争取阵地战中的和平所做的努力）。

我们现在假设在阶段 1 或者其他阶段，两家拍卖公司之前的收费标准都是 8%。此时就会有这个策略，一家拍卖公司的收费标准为 8%，为了求得心理平衡，这一标准创造的收益至少要等于更低的收费标准所产生的收益。从而：

$$\text{当}\frac{5}{1-\delta}\geqslant 7+\delta\left(\frac{4}{1-\delta}\right)\text{时，8\% 优于 6\%} \qquad (14.1)$$

$$\text{当}\frac{5}{1-\delta}\geqslant 4+\delta\left(\frac{4}{1-\delta}\right)\text{时，8\% 优于 4\%} \qquad (14.2)$$

上述不等式表示收取 8% 的费用时所获得的收益。这两个不等式成立的条件是其他公司当前的收费标准为 8%，从而所获的当前收益为 5。同时，两家拍卖公司以后的收费标准依然为 8%（此决定受各自战略的影响）。所得收益的现值是 $5/(1-\delta)$。不等式（14.1）表示收费标准定为 6% 时，拍卖公司所获得的收益。6% 的收费产生的当前收益为 7，而在确定的阶段的收益仅为 4，这一异常的结果使得两家拍卖公司都将收费标准定为 6%（取决于公司的策略）。如果两家公司下一时期都将收费标准定为 6%，那么依据他们的策略，他们以后的收费标准依然为 6%。因此，当前的收费标准不是 8% 时所产生的未来收益为 4，把这一收益折算成现值，表示成 $\delta[4/(1-\delta)]$。

据不等式（14.1）可知，当 8% 的收费标准所产生的收益大于等于 6% 的收费标准所产生的收益时，纳什均衡成立。我们用相似的思维分析，8% 的收费标准定然优于 4%，因此不等式（14.2）是正确的。背叛合作伙伴将收费标准定为 4% 或 6% 唯一不同的结果在于所产生的当前收益；因为惩罚措施的存在，所产生的将来的所得收益是相同的——与将收费标准永远定为 6% 所产生的收益是相同的。

为了达到均衡，我们需要（14.1）和（14.2）两个不

等式。注意到不等式（14.2）右侧的表达式小于（14.1）右侧的表达式，也就是说如果存在一个条件使得不等式（14.1）成立，那么这个条件一定满足不等式（14.2）。换句话说，如果一家拍卖公司打算违背双方之间的定价协议，那么他应该把收费标准定为 6%。因此，我们只需保障不等式（14.1）成立。解不等式（14.1）得出 δ：

$$(1-\delta)\times\left(\frac{5}{1-\delta}\right)\geqslant(1-\delta)\times\left[7+\delta\left(\frac{4}{1-\delta}\right)\right]$$

$$\Rightarrow 5\geqslant 7(1-\delta)+\delta 4\Rightarrow \delta\geqslant\frac{2}{3}$$

当 $\delta\geqslant\frac{2}{3}$ 时，每家拍卖公司都想将收费标准定为 8%，他们也希望其他公司都定为 8%。

有人认为这一对策略是子博弈精炼纳什均衡，为了完成这一讨论，我们需要假设一个阶段和一段历史，即在前面的阶段，一家公司的收费标准为 8% 或者两家公司的收费标准都不是 8%。我们注意到每个策略都指引拍卖公司选择 6% 且所产生的收益为 $4/(1-\delta)$，或者表示成 $4+\delta\ [4/(1-\delta)]$。如果一家公司的收费标准为 8%，则他的当前收益较低，为 1，然而他未来的所得收益依然为 4。每家公司都希望另一家公司的收费标准为 6%，因此后者是正确的，并且根据他们的策略，两家公司未来都会将收费标准定为 6%。把这一逻辑应用于将来，因此 8% 的收费标准产生的现实收益为 $1+\delta\ [4/(1-\delta)]$，这一结果明显低于 6% 的收费标准所产生的收益。类似的，4% 的收费标准所产生的现实收益为 $2+\delta\ [4/(1-\delta)]$，这一结果要更糟。从而，**源于策略的决策才是最佳的**。因此，惩罚本身就是一种均衡。总而言之，当且仅当 $\delta\geqslant\frac{2}{3}$ 时，这一对策略是一个子博弈精炼纳什均衡。

14.1.2 暂时适度的收费标准

假设阶段 1，两家拍卖公司的收费标准都很高，为 8%，且只要双方保持一致，他们未来的收费标准依然为 8%。如果他们都认为对方的收费标准是 8%，而其中有一家不是 8% 或者两家都不是，那么此时就存在不诚信行为，假设他们在三个阶段将收费标准定为 6%，过后重新定为 8%。如果按照规定，双方都应将收费标准定为 8%，而其中的一家或两家没有做到，那么他将受到惩罚，被要求将收费标准降为 6% 且维持三个阶段。实际上，每次不诚信过后都将受到惩罚，即连续三个阶段将收费标准降到 6%。

设想一段历史，假设两家拍卖公司现在的收费标准都为 8%。阶段 1 或者其他一个阶段，拍卖公司的收费标准为 8%，或者在一个阶段，他们结束了惩罚将收费标准重新定为 8%，8% 的收费标准所产生的收益为：

$$5+\delta 5+\delta^2 5+\delta^3 5+\cdots \quad (14.3)$$

这一结果必须大于或等于任何其他收费标准所产生的收益。接着我们有足够的数据推算最佳收费标准所产生的收益。一旦违背协议就会受到惩罚，而不管收费标准降为 4% 还是 6%，所受到的惩罚是一样的，因此最佳收费标准应为 6%，因为它所产生的当前收益较高。不诚信产生的收益为：

$$7+\delta\times 4+\delta^2\times 4+\delta^3\times 4+\delta^4\times 5+\delta^5\times 5+\cdots \quad (14.4)$$

我们注意到在必须接受惩罚的三个阶段，背叛协议产生了低收益 4，接着将来所获得的收益都为 5。这两个表达式显示从阶段 5 开始之后的收益是相同的，因此（14.3）大于或等于表达式（14.4）时均衡成立。我们可以忽略阶段 5 之后的那些时期，故而均衡条件为：

$$5+\delta\times 5+\delta^2\times 5+\delta^3\times 5\geqslant 7+\delta\times 4+\delta^2\times 4+\delta^3\times 4$$

$$\Rightarrow \delta+\delta^2+\delta^3\geqslant 2 \quad (14.5)$$

当然，我们需要确认这一惩罚措施是否是最佳的，既然这一惩罚包含三个阶段博弈的纳什均衡，这就证明了这一论证的正确性（因为实践证明每家拍卖公司都努力遵守协议避免惩罚才是最佳策略）。

解不等式（14.5），得出结果接近于 $\delta\geqslant 0.81$。（相信我，除非你的结果不是通过解立方等式得到的。）折现系数的制约因素比之前我们所述的更加严格。当惩罚要求其收费标准始终为 6% 时，拍卖公司只需保持足够的耐心，反应到折现系数中至少为 0.67。现在，我们仅用为期三个月的惩罚阻止所有的拍卖公司降低其收费标准，折扣因素至少为 0.81。通过减弱惩罚的力度——违规行为意味着仅在三个周期获得较低收益 4 而不是永远，为了遏制当前受利益诱惑违背协议的不诚信行为，参与者必须更多地考虑将来收益。


点拨

更为严格的惩罚措施有利于协调参与者之间的行为，更好地维持合作关系。一旦采用这一严厉的惩罚措施就意味着较低的收益，然而这一严格惩罚措施的震慑力更有利于维持合作关系。

14.1.3 暂时保持较低的收费率

"既没有把价格定得低于成本，也没有股东大声疾呼要求降低价格，只有这样的公司才会招来更多报复，更加悲哀。竞争是地狱。如果上将威廉•特库姆塞•谢尔曼（William Tecumseh Sherman）经营一家航空公司，他或许会这么说。

在一些工业领域，一旦相互间的统一定价机制瓦解，价格就会降低，甚至要低于正常竞争机制下的价格。实际上，他们甚至会以低于成本的价格销售商品，此时公司会出现亏损。一般而言，如此低的价格不会长久维持，然而，**短暂的低价倾销就是我们所熟知的价格战。**

有一个策略包含了价格战的思想，此战略中的惩罚措施比单纯回到阶段博弈的纳什均衡要严厉得多：

- 阶段 1，收费标准为 8%。
- 其他任何阶段：
  如果满足以下两个条件中的一个：（1）两家拍卖公司之前都将收费标准定为 8%；（2）之前都定为 4%，则将收费标准定为 8%；如果不满足上述条件则定为 4%。

首先注意到，如果按这一策略行事，则两家拍卖公司都会将收费标准定为 8%。为了阻止拍卖公司将收费标准降到 4%，我们制定了惩罚措施：除非两家公司都降到 4%，否则擅自降价会受到惩罚，他们将被要求保持 8% 的高收费标准。

与这一策略类似的是子博弈精炼纳什均衡，为了得到使这一命题成立的条件，我们首先假设，无论是阶段 1 还是前期的某一阶段，两家拍卖公司的收费标准不是 8% 就是 4%。他们一贯的选择 8%，且在所有阶段产生的收益为 5。这一收益大于或等于收费标准为 4% 或 6% 时所产生的收益。这两者中的最优选择为 6%，其产生的收益为：

$$7+\delta\times 0+\delta^2\times 5+\delta^3\times 5+\cdots$$

违背协议的公司获得的当前收益为 7，接下来他获得的收益为 0，因为这两家公司的收费标准都为 4%。然后他们将收费标准定为 8%，所产生的未来收益为 5。均衡条件是：

$$5+\delta\times 5+\delta^2\times 5+\delta^3\times 5+\cdots \geqslant 7+\delta\times 0+\delta^2\times 5+\delta^3\times 5+\cdots$$

$$\Rightarrow 5+\delta\times 5\geqslant 7\Rightarrow \delta\geqslant\frac{2}{5} \tag{14.6}$$

为了完成这一分析，我们需要确保惩罚的均衡性，即我们确信迫于惩罚机制的震慑力，两家拍卖公司将努力遵守协议。假设在前期，两家公司的收费标准既不是 8% 也不是 4%。那么按照他们惯常的行为，将会把收费标准定为 4%，所产生的收益为：

$$0+\delta\times 5+\delta^2\times 5+\delta^3\times 5+\cdots \tag{14.7}$$

假设另一家公司的收费标准也是 4%，此时会出现零收益，但是接着会升至 5，因为一旦价格战结束，两家公司会将收费标准重新定为 8%。不管收费标准是 6%，还是 8%，所产生的未来收益是相同的，因此，相对于 4%，6% 是最佳选择，因为 6% 产生的当前收益更高，也就是 1（8% 的收费标准产生的收益为-1）。6% 的收费标准产生的收益为：

$$1+\delta\times 0+\delta^2\times 5+\delta^3\times 5+\cdots \tag{14.8}$$

公司在现阶段所获得的收益为 1，但是根据其策略，下一阶段他将卷入价格战，价格战结束后，收费标准再次变为 8%。只要表达式（14.7）大于或等于（14.8），一旦有可能开始价格战，两家公司都会积极参与。即：

$$0+\delta\times 5+\delta^2\times 5+\delta^3\times 5+\cdots \geqslant 1+\delta\times 0+\delta^2\times 5+\delta^3\times 5+\cdots$$

$$\Rightarrow \delta\times 5\geqslant 1\Rightarrow \delta\geqslant\frac{1}{5} \tag{14.9}$$

总而言之，当表达式（14.6）和（14.9）同时成立时，这一对策略是子博弈精炼纳什均衡。表达式（14.6）确保一家拍卖公司希望维持合作关系（将收费标准定为 8% 且希望另一家公司的收费标准也是 8%）；表达式（14.9）确保出于形式所需，一家拍卖公司愿意参与存在竞争机制的价格战。当 $\delta\geqslant\frac{2}{5}$ 时，两个条件都成立。

这里有两个条件很重要。**首先，惩罚需要其专属合作的保障。**在价格战中，将收费标准定为 6% 比 4% 更有利于提高当前收益。8% 的高收费标准会产生更多的好处，基于这一诱惑，拍卖公司愿意参与备受指责的价格战。只有当公司定的收费标准较低且获得的收益较低时，他们才愿意订立较高的收费标准以获取较高的收益。为了获得最好的收益，拍卖公司需要通过改革快速恢复正常。当然，当表达式（14.9）的条件成立时，为了得以参与价格战，公司必须充分考虑其未来收益。

其次，"一旦违反规则要永远维持正常的竞争环境"这一惩罚措施的震慑力高于一个阶段价格战的震慑力，

因此后者更有利于价格垄断行为的发生。当 $\delta \geqslant \frac{2}{5}$ 时，价格战这一惩罚可以使统一定价关系更加稳固；然而，当 $\delta \geqslant \frac{2}{3}$ 时，阶段博弈的纳什均衡将永远成立。折现系数不必太高时拍卖公司更容易进行统一定价，换句话说，拍卖公司不必太多考虑未来收益。相比“长期而温和”的惩罚，一个“野蛮而短命”的惩罚更有效。

14.1 理解测验

以克里斯蒂公司和苏富比公司的无限次重复博弈为例。假设惩罚是二阶段的价格战，期间两家拍卖公司的收费标准降为4%。以此构建一对子博弈精炼纳什均衡。

14.1.4 更公正的惩罚

以牙还牙，以眼还眼。

——圣经（出自 21:23–27）

我们已经讨论过多种惩罚：永远保持竞争，暂时维持竞争关系，开始短暂而激烈的价格战。然而这些惩罚伤害了所有参与者的利益。即两个参与者，一方背信弃义另一方必会利益受损。较为公正的惩罚只是惩罚违背合约规定的参与者。我们现在开始讨论这一惩罚机制。

合作使得两家拍卖公司每一阶段的收费标准都是8%。如果克里斯蒂公司违背合作原则将收费标准降到8%以下，而苏富比公司依然保持8%，惩罚将导致苏富比公司与克里斯蒂公司遭受相同的损失。即下一阶段，克里斯蒂公司的收费标准为8%，而苏富比公司将收费标准降为6%。如果克里斯蒂公司确实能保证始终保持8%的收费标准，那么在随后的阶段，两家公司都会选择8%。基于惩罚机制，不诚信的参与者被要求选择较高的收费标准并且允许竞争对手降低收费标准以扩大业务。这种现象只会维系一个阶段，接下来他们会继续维持合作关系。如果未来有可能出现违约行为，我们必须确保相同的惩罚机制对其进行约束。最后，如果偶然的，这两家公司同时违约，此时，这种行为将免于惩罚，两家公司继续维持合作关系。

假设两家拍卖公司通过合作协商将收费标准定为8%，则当：

$$5+\delta\times 5+\delta^2\times 5+\cdots \geqslant 7+\delta\times 1+\delta^2\times 5+\cdots$$

$$\Rightarrow 5+\delta\times 5 \geqslant 7+\delta \Rightarrow \delta \geqslant \frac{1}{2}$$

时，此决策是明智的。收费标准为8%时，所获得的收益永远为5，违背协议将收费标准定为6%时所产生的收益最高（现实收益为7）会接着招致惩罚。后者包含下一阶段的收益1（因为一家公司违背协议将收费标准定为8%而另一家公司的收费标准为6%）接着双方都定为8%。

允许竞争对手降低价格是对违约者的惩罚，我们必须确保这种惩罚是最有效力的。因此，假设苏富比公司前期有不诚信行为，那么现阶段当克里斯蒂公司的收费标准为6%时，苏富比公司要定为8%，其所获得的当前收益为1。苏富比公司的最佳选择是与克里斯蒂公司保持一致将收费标准定为6%，此时按其策略在确定阶段努力贯彻协约避免惩罚。因此，当：

$$1+\delta\times 5+\delta^2\times 5+\cdots \geqslant 4+\delta\times 1+\delta^2\times 5+\cdots$$

$$\Rightarrow 1+\delta\times 5 \geqslant 4+\delta \Rightarrow \delta \geqslant \frac{3}{4}$$

时，苏富比公司的最佳收费标准是8%，且允许其竞争对手克里斯蒂公司降低收费标准。

我们注意到被欺骗的公司很乐意执行此惩罚措施。因为当违约公司将收费标准定为8%时，未违约公司可以将收费标准定为6%，所获得的当前收益为7。既然无论采取什么策略，未来所获得的收益是相同的，故而他想最大化当前收益，将收费标准定为6%。

惩罚就是以其人之道还治其人之身。如果苏富比公司没有按照合约规定将收费标准定为8%而是降低收费以扩大业务，那么一旦重新建立合作关系，克里斯蒂公司定会有机会对其进行报复。这就是亘古不变的因果报应。然而，违约会得到好处，即违约获得的收益高于遵守协议所获得的收益。

14.2 礼尚往来

“礼尚往来”不仅是人类的处事原则，动物王国里的其他物种也在努力恪守。拉丁语中的“quid pro quo”意为“交换”。许多时候这种交换几乎是同时发生的（比如黑猩猩轮流给对方梳理体毛）。然而，很多时候付出和回报并不是同步的。实际上，虽然你帮助过别人，但除非你同样需要帮助，否则你的付出将得不到相应的回报。

我们将通过两种吸血物种（吸血蝙蝠和政客）的活

动去探索此种形式的合作——相互帮助。第一个案例告诉我们："礼尚往来"如何导致美国国会挥霍浪费现象的激增。参议院的议员们相互支持彼此喜欢的项目（由纳税者承担费用），以使其在下一轮选举中得到帮助。类似的，吸血蝙蝠在必要的时候相互帮助，不过它得到的回报不是选票，而是同伴提供的维持其生命所需的血液——肮脏粗鲁得令人作呕。可是纳税人所付出的远不止这些。

◯案例研究

美国国会和建议经费的花销计划

在美国历史上受美国国会控制的铁路是最昂贵的公共交通工具，它为计划经费开支创造了一个新的纪录，花费 2 864 亿美元修建一条铁路。240 亿美元的专项资金用于 6 376 个项目，几乎每个地区的国会都这样制定经费计划。对于这一计划，众议院以 412:8 通过，参议院以 91:4 通过。

——杰夫·雅各布（Jeff Jacoby），《波士顿环球报》

给予政府金钱和权力就像把威士忌和汽车钥匙交给未成年男孩。

——P.J. 奥鲁尔克（P.J.O'Rourke），《议会娘子》(Parliament of Whores)

2005 年美国国会提出"共享便捷的交通"计划并拿出一笔资金，即从其他项目中挪用 22.3 亿美元建一座桥，这座桥连接凯奇坎（Ketchikan）和阿拉斯加这两座城市，而阿拉斯加的人口不足 9 000 人，其航空港建在格拉维纳岛上（Island of Gravina, 约 50 人）。尽管此计划确实将惠及阿拉斯加的几千人（使用桥的人及被雇用修桥的人），但花费如此巨额资金惠及区区几千人，许多人认为此项目荒唐无比。人们戏称这座桥为"没有终点的桥"，它的曝光引发了激烈的抗议。

修建没有价值的桥是经费计划中的一个经典案例。经费计划是为了惠及参选官员而提出的，尽管缺乏正当的理由，其存在确实有助于提高参选官员下一轮选举中的中选概率。这个战略旨在从巨大的人口基数中获取资源：在凯奇坎建桥案例中，联邦政府的所有纳税人都要为桥的修建承担费用，而此项目的好处集中在少数人身上。这是参议员泰德·史蒂文斯（Ted stevens）试图通过惠及其家乡的居民进而有利于其参选的一种方法。

当然，也许有人会问，为什么美国国会会提出经费计划这一项目呢？国会成员间的工作是对"礼尚往来"这一格言的最好验证。一旦众议员或参议员明白这一项目会影响其参选结果时，他们会把自己的经费计划放在最重要的位置。这样，他的资助项目会成为其参选最好的理由。这种形式的合作惠及了美国国会成员，每个议员都把资金用于有利于其选区居民的项目建设，而这笔资金来源于全国人民。

为了探究"礼尚往来"经费计划的稳定性，假设美国参议院仅有三个议员：参议员布朗、伯德和史蒂文斯。每个时期，他们中的一个人会提出一项经费计划。如果资金申请通过，项目发起人所获收益为 100，而其他两个成员所获收益为 25。假设资金申请未通过，所有成员都将获得零收益。我们注意到当国会议员看到其他成员所获得的收益时他会更不愉快，因为他的选区没有得到任何好处却必须为此承担费用。尽管这与正常的参选程序相反，但如果我们假设三个参选者只有申请到经费才有资格成为候选人，那么将大大简化我们的分析过程。假设经费计划是以一种交替的方式被提出的，即每个参议员每三个阶段提出一个计划。参议员布朗按照 1,4,7…这样的规律提出计划，伯德按 2,5,8…这样的规律提出计划，史蒂文斯按 3,6,9…这样的规律提出计划。

每一时期会有一个项目受到重视，而其他参议员必须决定是否支持这一项目。假设，起初所有的参议员赞成每一个项目。接着一个参议员不守信誉反对某议员的项目，这样其他两个议员会反对背叛者的项目。一旦这种假设成立，他们会再次相互支持所有的项目。而这正如前面几章所提到的，是一个不公正的惩罚。

假设当前阶段，所有的参议员紧密合作，相互支持彼此的项目。如果现在是参议员布朗提出项目的时候。当然，对他来说，支持自己的项目是最明智的选择，问题在于参议员伯德和史蒂文斯是否会赞成。如果参议员伯德赞成此项目，他获得的收益为：

$$-25+\delta\times100+\delta^2\times(-25)+\delta^3\times(-25)+\delta^4\times(100)+\delta^5\times(-25)+\cdots \tag{14.10}$$

如果参议员布朗的经费申请得到批准，那么伯德将得到-25 的收益，但是一旦他的项目得到批准，他在下一阶段所获得的收益为 100，而如果史蒂文斯的项目得

以通过，他在那一阶段将获得−25 的收益。这一收益结果是循环往复的。或者说，参议员伯德反对布朗的项目，此时，参议员伯德的收益为：

$$0+\delta\times0+\delta^2\times(-25)+\delta^3\times(-25)+\delta^4\times100+\delta^5\times(-25)+\cdots \quad (14.11)$$

当前这一阶段，因为参议员布朗的项目被否决，所以伯德所获收益为零。下一阶段，因为其他人为了报复否决了伯德的项目，他所获得的收益也为零。之后，为了彼此的利益，他们会重新开始合作，相互支持彼此的项目。

当表达式（14.10）大于等于表达式（14.11）时，对参议员伯德来说赞成此项目是最好的选择。据上述式子推出从第三个阶段开始所获得的收益是相同的。取消那些收益相同的阶段，我们得知当 $-25+\delta\times100\geqslant0\Rightarrow\delta\geqslant0.25$ 时，表达式（14.10）大于等于表达式（14.11）。

我们继续假设参议员布朗的项目受到重视，现在考虑参议员史蒂文斯。伯德和史蒂文斯所处境遇的不同之处在于史蒂文斯有两个阶段的低潮而不是一个阶段。

$$-25+\delta\times(-25)+\delta^2\times100+\delta^3\times(-25)+\delta^4\times(-25)+\delta^5\times100+\cdots\geqslant0+\delta\times0+\delta^2\times0+\delta^3\times(-25)+\delta^4\times(-25)+\delta^5\times100+\cdots \quad (14.12)$$

当上面的等式成立时，选择赞成是最佳策略。因为不管他对一个项目还是两个项目投反对票，其所受的惩罚是相同的。如果参议员史蒂文斯反对布朗的项目，那么他应该也会反对伯德的项目。从而，参议员史蒂文斯在第一阶段和第二阶段获得零收益，接着在第三阶段，因为其他参与者为了报复否决了他的项目，他依然获得零收益。之后，他们重新开始合作，坚持“礼尚往来”。从第四阶段开始，（14.12）两边的表达式是相同的，因此可以被写成：$-25-25\times\delta+100\times\delta^2\geqslant0\Rightarrow\delta\geqslant0.64$

到参议员伯德提出项目时，对此的分析与上述案例是相似的。当 $\delta\geqslant0.25$ 时，对史蒂文斯来说，赞成此项目是最佳策略，而当 $\delta\geqslant0.64$ 时，参议员布朗投赞成票将是最佳选择。到史蒂文斯提出项目时，当 $\delta\geqslant0.25$ 时布朗赞成，$\delta\geqslant0.64$ 时伯德也会赞成。最后，如果这一案例需要一套惩罚机制，假设所有的人都反对其他人提出的项目，那么显然跟大家保持一致投反对票是明智的。

总而言之，当 $\delta\geqslant0.64$ 时，赞成其他人的赞助项目这一规则是均衡的（同时要保证 $\delta\geqslant0.25$）。如果下一阶段轮到其中的一个参议员提出项目，为了保证得到他人的赞同，折现系数要大于等于于 0.25。然而，如果连续两个阶段都轮不到他提出项目，那么他必须有足够的耐心。反映到折现系数上，要保证折现系数至少为 0.64。因为惩罚是未来才会到来的，为了更好地贯彻所提出的项目，参议员必须更多地考虑未来收益。

吸血蝙蝠及互利主义

叶口蝠，也就是我们常说的吸血蝙蝠，其生存环境非常恶劣。吸血蝙蝠通过吸食动物（牛和马）而不是人类的血液维持生存。吸血蝙蝠每晚的吸血量相当于其体重的 50%~100%。如果没有足够的血液供应，它们存活不了很久，如果几天没有吸到血，它们的体重会下降 25% 并且其体温会降到非常低甚至危及生命。

吸血蝙蝠靠频繁的血液消耗维系生命，它们之间建立了一种合作关系，生物学上称之为**互利主义**（reciprocal altruism），即一只吸血蝙蝠最近没有进食且急需食物，此时当晚成功进食的另一只吸血蝙蝠会把自己吸到的血给它。有趣的是，这种相互间的食物分享行为仅存在于母性吸血蝙蝠之间，尽管它们也给小蝙蝠喂食。我们暂且不去讨论这种行为是自然选择的结果还是后天形成的。现在用博弈论解释这种稳固的合作关系。

为了简化分析，假设一个群体中只有两只吸血蝙蝠。每天晚上它们飞出去寻找血液以维系生命，黎明的时候，一只吸血蝙蝠吸足了血或者空着肚子回来。我们假设吸血蝙蝠晚上独自出去觅食会有三种可能：第一，两只吸血蝙蝠都成功找到了血液，此时它们不需要分享；第二，它们都没有找到血液，此时没有血液可以分享；第三，一只吸血蝙蝠成功觅食而另一只没有找到血液，此时成功觅食的吸血蝙蝠要决定是否在每个找到血液的夜晚跟另一只蝙蝠分享食物（见表 14—1）。

表 14—1 吸血蝙蝠的收益

蝙蝠	分享	不分享
吃饱的	8	10
饿着的	4	−1

表 14—1 列出了吸血蝙蝠的收益取决于其中一只没有找到食物时分享行为是否发生。如果不存在分享，则成功觅食的吸血蝙蝠的收益为 10，而另一只的收益为−1。

相反，如果它们相互分享食物，因为相对于其找到的血液，成功觅食的吸血蝙蝠消耗的血液减少了，其所获得的收益为 8。但是觅食失败的吸血蝙蝠的收益升到了 4。

分析重复博弈时我们使用了假设法，假设每只吸血蝙蝠都最大化自己的收益。现在，当你设想非人类忙于算计的时候你或许有一点儿不舒服，但是实践表明，鸽子和老鼠确实更看重当前收益而不是未来收益（尽管据我所知没有人对吸血蝙蝠做过这种实验）。折现系数也可以解释成活到下一阶段的概率。当你发现自然选择导致对动物来说，今天的食物比明天的食物更重要时，你就不再惊讶于这一现象了，因为没有食物它或许活不到下一阶段。我们假设吸血蝙蝠是精于算计的，如果你继续为这一假设感到不舒服，我们用真实的吸血鬼系列电视节目《捉鬼者巴菲》(*Buffy the Vampire Slayer*）中的安雅（Anya）和斯柏克（Spike）代替吸血蝙蝠。

吸血蝙蝠的策略告诉它，一只吸血蝙蝠没有找到食物时，成功觅食的吸血蝙蝠是否与其分享取决于它们间的合作史。假设成功觅食的吸血蝙蝠一直与找不到食物的同伴分享自己的食物，那么此时它还会与其共享食物。相反，如果它们从未相互分享过食物，此时也不可能相互分享。这就是所谓的严峻的触发战略。让我们得出这一条件使其成为子博弈精炼纳什均衡。

作为我们分析的第一步，吸血蝙蝠出去觅食之前，也就是说吸血蝙蝠尚不知是否能找到食物之前，我们将通过这一策略得知期望收益。如果吸血蝙蝠乐于相互分享。我们用 V 表示期望收益，得出等式：

$$V=s\times[s\times10+(1-s)\times8]+(1-s)\times[s\times4+(1-s)\times(-1)]+\delta V \qquad (14.13)$$

我们解释一下这一等式。假设一只吸血蝙蝠成功觅食并独享食物的概率为 s，当前阶段的期望收益为 $s\times10+(1-s)\times8$；另一只吸血蝙蝠成功觅食的概率也为 s，此时不存在分享，两只吸血蝙蝠的收益为 10；假设其中一只没有找到食物的概率为 $1-s$，此时分享导致成功觅食的吸血蝙蝠的当前收益减少至 8。这样我们得出了式子（14.13）的一部分：$s\times[s\times10+(1-s)\times8]$。还有一种可能，我们以找不到食物的吸血蝙蝠作为分析对象，即这只蝙蝠找不到食物的概率为 $1-s$，当其同伴成功觅食并与之分享食物时，它的收益为 4；当其同伴同样没有食物时，它的收益为−1。基于这一分析得出等式 $(1-s)\times[s\times4+(1-s)\times(-1)]$。（14.13）中上述两部分的和就是吸血蝙蝠的当前期望收益。第三项 δV 为未来期望收益。吸血蝙蝠下一阶段的期望收益为 V，而这一收益的折扣为 δ。现在我们简化关于 V 的最终表达，得出等式：

$$V-\delta V=10s^2+8s(1-s)+4s(1-s)-(1-s)^2$$

$$\Rightarrow V(1-\delta)=-3s^2+14s-1\Rightarrow V=\frac{-3s^2+14s-1}{1-\delta}$$

如果两只吸血蝙蝠之间存在背叛行为，既存的惩罚机制会让它们不再与同伴分享食物。假设一个阶段不发生共享行为，此阶段的期望收益为：

$$s\times10+(1-s)\times(-1)=11s-1$$

假设一只吸血蝙蝠成功觅食的概率为 s，其收益为 10；如果找不到食物的概率为 $1-s$，其收益为−1。因为这些假设收益适用于每一阶段，收益的现值为 $(11s-1)/(1-\delta)$。

我们现在可以确定成功觅食的吸血蝙蝠是否愿意与其同伴共享食物。假设两只吸血蝙蝠过去一直共享食物，此时当表达式（14.14）成立时，成功觅食的吸血蝙蝠与其同伴共享食物是最佳策略。

$$8+\delta\left(\frac{-3s^2+14s-1}{1-\delta}\right)\geqslant10+\delta\left(\frac{11s-1}{1-\delta}\right) \qquad (14.14)$$

通过分享，这只吸血蝙蝠的当前收益为 8，而当吸血蝙蝠预测到未来可以与同伴共享食物时，它的未来期望收益为 V。上述不等式的右侧是没有共享行为时的收益。现在假设两只吸血蝙蝠间不存在共享，它获得较高的当前收益 10，未来期望收益 $(11s-1)/(1-\delta)$。如果不等式（14.14）成立且它们间的共享行为从未中断时，共享将是最佳策略。

如果它们的共享史上发生过背叛行为，当：

$$10+\delta\left(\frac{11s-1}{1-\delta}\right)\geqslant8+\delta\left(\frac{11s-1}{1-\delta}\right)$$

时，成功觅食的吸血蝙蝠不与同伴分享食物实际上是最佳策略。

上述不等式是永远成立的。共享降低了当前收益，且没有较高的未来收益作为补偿。因为不管未来发生什么，其收益为 $(11s-1)/(1-\delta)$。因此，在没有遇见未来较高的补偿性收益时与同伴分享自己的食物是不明智的。

如果不等式（14.14）成立，这一互利主义策略就是子博弈精炼纳什均衡。经过一些简化运算，不等式（14.14）可以表示成：$s(1-s)\geqslant\frac{2(1-\delta)}{3\delta}$

为了方便下面的分析，假设 $\delta=0.8$，此时前提条件变为 $s(1-s)\geqslant\frac{1}{6}$。

请看图 14—2 中 s（1−s）所在的位置，当成功概率介于 0.21 到 0.79 之间时，互利主义是自我压制（也就是说，这是一种均衡），此时的收益既不会太高也不会太低。如果成功的概率太低，意味着当前阶段没有找到食物的吸血蝙蝠未来可以与同伴分享自己食物的机会太渺茫，不足以诱使现阶段成功觅食的同伴与自己分享食物。也就是说，现在的付出有可能未来得不到回报。相反，如果成功的概率太高，现阶段成功觅食的吸血蝙蝠不必担心未来没有食物，故而它独自进食的可能性很高。因为它未来似乎不需要同伴的帮助。最后我们得出这一结论，**当每个个体成功觅食的概率大体相同时，互利主义才会是稳固的。**

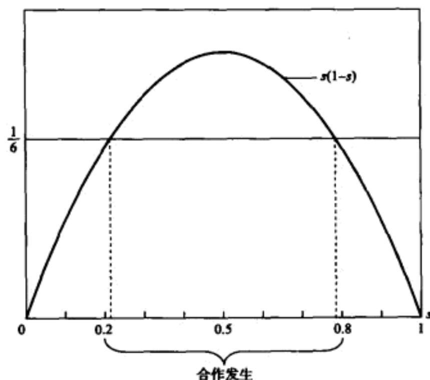


图 14—2 互利主义的均衡条件

14.2 理解测验

假设无限次重复博弈是以14.2小节中的阶段博弈为基础的。参与者1的策略让他偶尔选择a，大多时候选择c，选择a时的结果为（a,z），选择c时的结果为（c,x）。其他阶段，参与者1选择b。参与者2的策略让他偶尔选择z，大多时候选择x，选择z时的结果为（a,z）选择x时的结果为（c,x）。其他阶段，参与者2选择y。我们发现这一策略成立的条件是子博弈精炼纳什均衡。

参与者 2

		x	y	z
参与者 1	a	2,2	2,3	1,12
	b	3,2	4,4	3,1
	c	12,1	1,3	1,1

图 14—3



点拨

参与者足够在意未来收益且未来需要帮助的概率适中时，互利主义是可持续的。

14.3 名誉

> 名誉是你可能拥有的最宝贵的财富。赞扬就像火焰，一旦你把它点燃，你可能很容易就占有了它。但是如果一旦你把它熄灭了，你会发现想重新点燃它是非常困难的。
>
> ——苏格拉底

名誉是宝贵的财富。正直的名誉（正如买卖双方在 eBay 获得较高的回评分数）会使其他人更愿意做你的合作伙伴；努力工作的名誉会帮助你得到一份好工作；按时还债的名誉会帮助你顺利拿到下一笔贷款。同时，正如苏格拉底所说的，名誉是脆弱的，唯一的一次说谎会让大家都觉得你是骗子。一次未还债务可能会给你带来糟糕的信誉风险。名誉的脆弱是不幸的，但是正是这种不幸让我们认识到，给他人正确的鼓励使其行为得体是非常必要的。下面两个章节，我们将验证这一观点。

○案例研究

做国王的债权人

资金短缺时（比如需要资金支撑一场战争时），近代前欧洲的国王会向私营放款人借债。因为当时征税尚未普及，借款对国王来说是很重要的资金来源。当然有时候国王会赖账（比如 1290 年，国王爱德华一世将他的债权人犹太人驱逐出英国）。但是大多时候国王会按时还债。现在的问题在于国王是凌驾于法律之上的，为什么他还会偿还债务呢？

假设一个国王和一个唯一的放款人（你可以把这个放款人想象成一个统一的放款人组织，比如银行机构）。每个阶段的初期，国王需要借款的概率为“b”，他需要资金作战或购买土地。此时，放款人必须决定是否借款给国王。假设债权人收取 10% 的利息，也就是说，如果借出 100 单位的资金，按照协议国王要偿还 110 单位（100 这个基数加上 10 个单位的利息）。借款到期的时候，

国王将决定是否偿还这110个单位的债务。假设对国王来说这笔资金的价值是125，那么国王更愿意得到这笔借款，并按规定偿还而不是放弃借款。最后，假设放款人不借款给国王，他们的资金将得不到任何收益。因此，如果放款人认为国王会按期偿还债务，他会把钱借给国王。

假设国王偿还了过去的借款，只要国王需要，放款人会在第一时间提供资金并将持续这一交易。但是如果国王总是赖账，放款人将拒绝再次借钱给他。此时，如果国王过去经常按时还债，那么他的策略应该是偿还最初的和将来的借款。然而，如果他总是违约，那么他将拒绝偿还未来所有的借款。每个参与者都关注于如何使他的资金在当前获得最大的收益。

设想一下放款人的策略，如果国王有良好的信誉（他总是按规定偿还债务），对放款人来说再次借款给国王是明智的选择。因为每次放款都会产生10个单位的收益（放款100个单位，利息为10%）。如果与之相反，国王拒绝偿还过去的一笔借款，那么根据国王的策略，他有可能拒绝偿还以后产生的所有债务。此时，放款人不愿借款给国王，这一策略是合乎情理的。

这个有趣的分析将焦点集中到了国王的行为上，如果他把自己变成了不守信誉的人，那么对他来说，不偿还借款实际上是明智的。因为根据放款人的策略，不管国王如何改变他都不会再借款给国王。因为偿还借款他得到的收益为−110，然而，不偿还借款得到的收益为0（他已经从到手的借款中得到125的收益）。

最后，我们假设国王有很高的信誉并且他有一笔尚未偿还的债务。如果国王偿还这笔债务（未来国王按其策略行事），他的收益为：

$$-110+\delta\times b\times 15+\delta^2\times b\times 15+\cdots=-110+\delta\left(\frac{b\times 15}{1-\delta}\right)$$

当前阶段，国王为偿还债务付出了110单位的资金，而下一时期他需要借款的概率为“b”。假设他已经偿还了过去的债务，他能得到这笔借款。这笔新的借款产生的收益为15（国王认为这笔借款将产生125个单位的价值，而我们需要从中去除所要偿还的110个单位的资金）。类似的，假设国王随后每个阶段需要借款的概率为“b”，或者假设国王拒绝偿还当前的借款，此时他的当前收益为0（他没有损失资金）。而将来任一阶段，放款人不会再继续借钱给他。因此，当不等式（14.15）成立时，偿还这笔借款是国王的最佳策略。

$$-110+\delta\left(\frac{b\times 15}{1-\delta}\right)\geqslant 0\Rightarrow\delta\geqslant\frac{110}{110+15b}\qquad(14.15)$$

国王偿还借款捍卫了他守信誉的美德，使得他以后能顺利拿到借款。并且放款人预计他有一套让国王维持名誉的激励机制。因此，尽管他们知道国王凌驾于法律之上，可以在任何时候拒绝偿还债务，他们还是借钱给国王。

为了得到一些对未来有利的启示，探索均衡条件是很有现实意义的。假设“b”接近于0，即国王不希望得到另一笔借款。此时，因为折现系数必须接近于1，不等式（14.15）成立的条件是非常严格的。如果国王预测他以后很可能不需要借款，因此他不可能偿还当前的借款。这样的话，借款者会想到国王将拒绝偿还债务，因此他拿不回原来借给国王的资金。国王偿还借款唯一的理由就是他想让大家相信自己是守信誉的，以便得到更多的借款。如果他认为下一笔借款将是最后一次。那么他就失去了偿还借款的动力。更为讽刺的是，如果放款人再也不希望国王做他的债务人，那么国王的信誉会非常差。

比起无足轻重的国民，国王的地位更为脆弱。因为放款人可以把国王以外的任何人送上法庭要求其偿还债务。我们有足够的理由相信每个人都会偿还借款。与之相比，国王偿还借款的可能性是值得怀疑的，因为没有一家法庭可以强迫国王偿还借款，并且国王自己也不能保证按规定偿还债务，他只是希望放款人相信偿还借款是对他最有利的选择以便赢得好的名誉。凌驾于法律之上的地位让国王借钱非同一般的困难。

福特公司及其日工资5美元的规定

福特汽车是世界上最大最成功的汽车制造业公司。1月12日它开创了一场汽车领域最大规模的变革以激励那些曾经闻名于工业领域的工人。

——《底特律日报》(*Detroit Journal*)，1914

1914年，福特公司的日工资为5美元，这在当时是前所未闻的。当时汽车领域最高工资是每天2.25美元。尽管我们可以单纯地理解为福特公司只是对他的工人很慷慨。然而事实上这一决策增加了福特公司的收益。下面我们进行分析。

如果一个工人受雇于福特汽车公司，他一定会面对

这样的抉择，即是努力工作还是偷懒。假设工人努力工作的货币成本是 1，偷懒的货币成本是 0，而工人得到的工资是 w，因此如果他偷懒，在那一阶段得到的收益是 w，而努力工作得到的收益是 w-1。工人的收益是一个独立阶段所获收益的现值，他的折现系数是 δ。

福特公司的收益是他从每个员工身上所得利益的现值，β 是折现系数。假设公司每天从偷懒员工身上得到的收益为 4 美元，从努力工作的员工身上得到的收益是 7 美元。假如福特公司给工人的日工资是 5 美元，如果工人偷懒他将被立即解雇，新员工会努力工作并将持续这种状态。此时，福特公司得到的收益为：

$$(4-5)+\beta(7-5)+\beta^2(7-5)+\cdots=-1+\beta\left(\frac{2}{1-\beta}\right)$$

假设市场上有充足的劳动力供应并且他们的期望日工资是 5 美元，同样，市场上也有充足的工作岗位。也就是说，即使一个工人被辞退了，他也能找到一个日工资 3 美元的职位且允许他不努力工作。在这一博弈中，福特公司要决定付给工人多少工资及是否解雇偷懒的员工。如果他解雇了一名员工，则必然要雇用一名新员工，这个新员工要决定是否为福特公司努力工作。

假设福特公司和他的员工签订了一条契约（作为公司的策略）：只要员工努力工作，他将得到 5 美元的日工资并将被长期雇用；一旦员工偷懒，他将被大家认为是偷懒的员工并将被公司辞退。因此，据员工的策略，日工资大于等于 5 美元时他们会努力工作，只要低于 5 美元他们就会偷懒。

要评价福特公司的策略是否为最佳策略，首先要明白福特公司对解雇员工还是留住员工没有多大的兴趣，因为他们随时可以找到新员工。对福特公司来说，解雇偷懒员工留住努力工作的员工将是最佳策略（这一策略能够让福特公司直接获得最高的收益）。真正的问题在于 5 美元的工资是否合理。假设公司每天给员工的工资为 w，并且未来会按照其策略行事（即支付给工人 5 美元的工资）。并且假设员工也按照他们自己的策略行事，则福特公司的收益是：

$$\begin{cases}(7-w)+\beta\left(\frac{2}{1-\beta}\right), & 5\leqslant w\\ (4-w)+\beta\left(\frac{2}{1-\beta}\right), & 3\leqslant w<5\\ \beta\left(\frac{2}{1-\beta}\right), & w<3\end{cases}$$

如果日工资大于等于 5，员工会努力工作，因此福特公司得到的当前收益为 7-w，并且如果他未来依然据其策略行事，得到的收益为 β [2/（1-β)]，因为他前一阶段的收益为 2(总收益 7，支付工资 5)。如果 $3\leqslant w<5$，员工就会偷懒，福特公司当前收益仅为 4-w。最后，如果工资低于 3，没有工人愿意受雇于福特公司。因为他们在任何公司都能得到 3 美元的工资且工作更轻松。这就意味着福特的当前收益为 0，尽管他的未来收益依然为 β [2/（1-β)]

很显然，福特公司不愿意支付高于 5 美元的工资，因为工人的期望工资为 5 美元，公司没必要支付高于 5 美元的工资激励员工努力工作。当然，如果在 3 美元和 5 美元之间选择，他更愿意支付 3 美元，因为高于 3 美元的工资依然不能激励员工努力工作。最后，相比任何低于 3 美元的工资，他更倾向于支付 3 美元。因为支付 3 美元的工资能产生 4 美元的收益，从而获得 1 美元的当前收益。然而工资一旦低于 3 美元，没有工人愿意来公司工作，从而产生零收益。因此对公司来说，最佳工资是 3 美元或 5 美元。据福特公司的策略：

$$2+\beta\left(\frac{2}{1-\beta}\right)\geqslant 1+\beta\left(\frac{2}{1-\beta}\right)\Rightarrow 2\geqslant 1$$

此时福特公司愿意支付 5 美元的工资。

换句话说，福特公司遵照其策略行事。如果一个员工努力工作创造的额外收益大于等于公司为了激励员工努力工作而增加的工资费用，那么福特公司的策略为最佳策略。

现在，我们假设有一个员工，福特公司的策略是既定的，并且这个员工未来会按照其既存的策略行事。据其策略，当：

$$\frac{5-1}{1-\delta}\geqslant 5+\delta\left(\frac{3}{1-\delta}\right)\Rightarrow\delta\geqslant\frac{1}{2}$$

时，他更愿意努力工作而不是偷懒。因此，只要员工的折现系数大于等于$\frac{1}{2}$，他更愿意为了 5 美元的工资努力工作而不是偷懒丢掉这个高薪职位。

这一策略成立的关键在于，对福特公司来说，员工努力工作创造的价值高于为了激励员工所花费的成本。通过员工努力工作，公司可以从每个员工身上得到 3 美元的额外收益，而他仅需要支付给员工 1 美元以确保他们努力工作。因此，员工努力工作时，福特公司每天得到 2 美元的净收益。让工人更加努力工作的方法是支付

给他们5美元的工资（高于其他公司的工资标准）。如果他们不努力工作就会被解雇。只要员工充分考虑了他的未来工资，他会努力不懈的工作以维持努力工作的名誉。根据这一策略，双方的经济都会更加宽裕，福特公司每天从每个工人身上多获得1美元的收益而员工每天多得到1美元的工资。

最后，有必要提醒我们自己，生活比我们想象的更加复杂。其实，就像福特公司支付给员工5美元的日工资这一案例，看起来公司不只是看重收益。除了慷慨的理念，公司还愿意改善其员工的生活水平。

为了有资格得到5美元的工资，雇员需要向公司表明他不喝酒，不对家人施暴，家里有寄宿生，他经常通过节省开支去存钱以保持一个整洁的家庭环境，以及一个较高的道德品质。

鉴于人类的行动力，他们不仅仅是宣传，实际上也是在创造。应该成立一个“社会学部门”监视他的员工（即使员工在家也被监视）。此时，员工就会怀疑为了5美元是否值得付出如此高的代价。

14.4 有缺陷的监视体系和反弹道导弹

> 美国和苏联认为核战争将给人类带来毁灭性的灾难。基于这一假设，双方认为，采取有效措施限制反弹道导弹系统（ABM）是限制战略军事进攻竞赛、降低核武器战争爆发概率的重要因素。就此，双方达成以下协议：双方政党承诺，不为了保卫本国领土而改动ABM系统。本着服从协议的意图，每个政党可以任意使用国家信息手段去证实其行为与国际法的准则相吻合。
>
> ——《限制反导弹系统协议》
>
> （*Treaty on the Limitation of Anti-Ballistic Missile Systems*）

因为不存在第三个政党强制两个国家履行彼此间的协议，国际条约只有在双方的自我约束下才能发挥作用，并且订立协约的双方必须在任何时候都愿意履约。我们可以把签订协约的两个国家比作签订合同的买卖双方，只要卖家交付了商品，买家就必须付款。如果买家不付款卖家可以将其告上法庭强制要求他付款。假如没有警察和法庭，买家几乎没有理会由会为已经到手的商品支付费用。

依靠自我约束力履行条约这一必要条件能够最大限度地保障国家间的条约得以订立并成功实施。国际条约的第二个问题在于如何监视缔约方的行为，一个国家可能怀疑对方未遵守条约，只是没有进入其他国家证实其怀疑的权力。2003年对伊拉克的包围入侵依然无法证实伊拉克是否藏有大规模杀伤性武器。这一事件证实监视机制无法保证条约履行。

让我们探讨一下如何在非完善的监视体系中维持合作关系。为此，我们准备用首字母表示一些名词。冷战期间，美国和苏联之间订立的最重要的条约之一是反导弹协议（ABM协议），这一协议源于限制军备会谈（SALT）。订立反导弹协议是为了击落导弹以保卫国家。最初担心ABM系统的存在更容易引发战争，但是防御性武器怎么会更容易导致战争呢？在此，我们有必要明确这一观点，**相互毁灭战略（MAD）的关键在于没有一个国家会首先发动攻击彻底摧毁其他国家的核武器。**也就意味着，如果美国进攻苏联也将给美国带去毁灭性的灾难，因为苏联同样也会向美国发射核武器。除非一个国家有ABM系统，否则MAD将失去稳定性。因为在行之有效的ABM系统保障下，每个国家都有理由相信他可以首先发动一次高效率的进攻并且能够防御其他国家残留的核武器。糟糕的是，按照一些博弈论的逻辑，如果没有ABM系统的国家预测到他马上会遭到进攻，那么首先发动进攻才是明智的。因为MAD，双方都选择不使用核武器作战，而ABM系统会打破这一均衡。正是这一恐惧使尼克松总统与勃列日涅夫（Leonid Brezhner）于1972年签订了ABM协议。直到2002年美国退出协议，这一协议一直在发挥效用。

每一阶段，两个国家同时决定所需要的ABM系统，正如图14—4所示，ABM系统有三个等级。这一阶段博弈中存在独特的纳什均衡，即两个国家都选择高水平的ABM系统。然而这样产生的结果并不会惠及两个国家。因为如果选择低等级的ABM系统，双方最好是退出协议，这将使他们的收益从3上升至8，或者最好没有ABM系统，此时收益上升至10。

苏联

		无ABM	低ABM	高ABM
美国	无ABM	10,10	6,12	0,18
	低ABM	12,6	8,8	2,14
	高ABM	18,0	14,2	3,3

图 14—4　ABM 协议博弈中的阶段博弈

假设这个策略无限次重复，但是与本章及前几章关于这一策略运用的案例相反，这个博弈的历史并非共同知识。也就是说，没有一个国家可以确保得知另一个国家过去的选择。因此，确定另一个国家是否有 ABM 系统会很困难。我们在表 14—2 中已列出了监视机制的不确定性。如果一个国家没有 ABM 系统，那么另一个国家绝对不会发现其 ABM 系统；如果他选择较低等级的 ABM 系统，那么他被发现导弹的概率为 10%；当拥有较高等级的 ABM 系统时，概率升至 50%。

表 14—2　ABM 条约博弈中的被监视到的概率

ABM的数量	检测到的概率
无	0
低	0.10
高	0.50

表 14—2 列出了不同等级的 ABM 系统被发现的概率，每一时期，国家要做出选择，决定使用哪一等级的 ABM 系统。我们假定有一个国家发现了 ABM 系统这一事实是共同知识，之后，每个国家都寻找方法以最大化他们期望的当前收益的现值。

设想一个相似的策略阻止所有的 ABM 系统：

- 阶段一，选择没有 ABM 系统。
- 其他阶段，

  如果过去没有国家被发现过 ABM 系统，则选择没有 ABM 系统。

  如果过去的某些阶段，一个国家被发现过 ABM 系统，则选择高等级的 ABM 系统。

如果未曾发现违背协议的行为，则不采用 ABM 系统所产生的收益为 10/（1−δ）。如果一个国家选择较低等级的 ABM 系统，那么他的期望收益为：

$$12+\delta\left[0.1\times\left(\frac{3}{1-\delta}\right)+0.9\times\left(\frac{10}{1-\delta}\right)\right] \qquad (14.16)$$

背叛行为被发现的概率为 0.1，一旦被发现，两个国家会把相当大的精力放到 ABM 系统上。但是仍然有 90% 的概率背叛行为不被发现。此时，依据其策略，两个国家都将选择不使用 ABM 系统。

其中的一个国家会选择高等级的 ABM 系统而不是先选择低等级的 ABM 系统而后再背叛协约。这样他的期望收益为：

$$18+\delta\left[0.5\times\left(\frac{3}{1-\delta}\right)+0.5\times\left(\frac{10}{1-\delta}\right)\right] \qquad (14.17)$$

当这样做所产生的收益 10/（1−δ）大于等于式子（14.16）和（14.17）时，遵守条约将是最佳策略。此时，必须保证下面的不等式成立：

$$\frac{10}{1-\delta}\geqslant 12+\delta\left[0.1\times\left(\frac{3}{1-\delta}\right)+0.9\times\left(\frac{10}{1-\delta}\right)\right]$$

$$\Rightarrow \delta\geqslant\frac{2}{2.7}\approx 0.74 \qquad (14.18)$$

$$\frac{10}{1-\delta}\geqslant 18+\delta\left[0.5\times\left(\frac{3}{1-\delta}\right)+0.5\times\left(\frac{10}{1-\delta}\right)\right]$$

$$\Rightarrow \delta\geqslant\frac{8}{11.5}\approx 0.70 \qquad (14.19)$$

从而得出 δ 大于等于 0.74。如果 $\delta < 0.70$，最好的策略是选择不使用 ABM 系统。如果 $0.70 < \delta < 0.74$，低等级的 ABM 系统更有利于完成禁令。只有当 $\delta \geqslant 0.74$ 时，对每个国家来说，遵守 ABM 协议才是他们的最佳选择。

14.3 理解测验

假设科技的进步提高了监视水平，此时监测到ABM系统的概率如表14—3所示。我们使用刚才描述的策略收益得出均衡条件。如果你回答正确，你将会发现，折现系数的限制并不是很严格，也就表明较好的监视系统更有利于维持合作关系。

表 14—3　ABM 条约中被较高监测到的概率

ABM的数量	监测到的ABM概率
无	0
低	0.30
高	0.75

本章总结与练习题

本章总结

人类在个人利益的驱使下各自为战，结果会令每个参与者失望。相反，如果人类能够相互协调，分工合作，他们会得到满意的结果。就像本章作为例证所举的那些事例，在现实生活中也并不少见。这些事例可以存在于国家、国王、公司、政客、工人甚至吸血鬼之间。综合所有的事例我们得出这一结论，即重复博弈理论旨在告诉我们参与者在什么时候以怎样的方式创建一个机制以支持互惠行为。

这一机制的具体表现形式可能有所差异，但是参与者为了彼此的奖惩而合作的原则是始终不变的。如果每个参与者都遵守合作机制，作为回报，其他参与者将继续遵守合作原则。一些人背离了合作原则，作为惩罚，整个团队将对他发起进攻，如在艺术品拍卖公司这一事例中，强迫违背原则的公司降低收费率。当然，惩罚也会因违约者的地位、违约行为的严重程度及违约时间的长短而有不同的表现形式。有一种惩罚仅仅是终止合作，之后开始一个阶段博弈的纳什均衡。这种惩罚机制会让参与者在一个具体的时间，比如5个阶段（或者永远）回到各自为战的状态。而且这种状态持续的时间越长所受到的惩罚也越严厉，并且惩罚所产生的威慑力将更能有效地防止背叛行为的发生。为了使惩罚更具威慑力，有必要把收益降到阶段博弈纳什均衡所产生的收益之下。然而，如此极端的惩罚机制不可能永远有效地发挥作用。为了使参与者在如此严苛的惩罚机制的震慑下进行博弈，必须保证将来能重新开始合作。这种短期的严厉的惩罚已在市场上的价格战中得以体现，即回到盈利状态之前价格甚至能降到成本以下。

参与者过去的行为影响着他的**名誉**，大家会据他过去的行为判断他未来的表现，这也是一个了解如何维系合作关系的方法。如果一个参与者违背了合作原则，其他参与者会认为他不守规则，并且认为他以后也不会守信用，这一预测使他们改变自己的决策（有谁愿意跟一个可能背叛自己的同伴合作呢？），从而合作关系破裂。重复博弈似乎让我们明白了参与者在以何种方式维护自己的名誉。研究发现，国王通过偿还借款维护他“守信用”的名誉以诱使放款人继续借款给他。或者，雇员努力工作以赢得“先进员工”的荣誉称号，一旦他工作偷懒将得到“工作不积极”的评价并且将被辞退。

生活中有许多重复博弈，我们同我们的父母、兄弟姐妹、配偶、朋友、邻居、老板、同事等进行着博弈。许多情况下，本章及前几章总结出的支持合作机制得以运行的方法在偷偷地发挥作用。我们的关系在“我们会公平地对待彼此”这一理念的支持下得以建立。如果我们不能做到公平，可能会产生某种结果，尽管这一结果或许令人不快，它确实有着积极的作用，因为这些结果所产生的威慑力使人们以一种积极得体的方式进行博弈，这是每个参与者乐意看到的。

练习题

1. 设想一个无限次重复博弈，阶段博弈如图14—5所示。每个参与者的收益就是他所得收益的现值，折现系数为δ。

a. 确定一个冷酷策略，它导致参与者1选择c，参与者2选择y，并且这一策略的既定条件产生了一个子博弈精炼纳什均衡。

b. 考虑以下策略所产生的收益：阶段1，参与者1选择c。其他阶段，比如之前的阶段，如果结果为（c,y）或（d,z），参与者1会选择c，否则他会选择d。阶段1，参与者2选择y。其他阶段，如果之前阶段的结果为（c,y）或（d,z）参与者2选择y，否则他会选择z。为得到这一收益总结出满足条件以便得到一个满足纳什均衡

的子博弈。

参与者 2

		w	x	y	z
参与者 1	a	2,2	3,1	2,0	4,−4
	b	1,3	4,4	3,1	2, 3
	c	0,2	1,3	7,7	3, 9
	d	−4,4	3,2	9,3	0, 0

图 14—5

c. 考虑如下策略及产生的收益：阶段 1，参与者 1 选择 c。其他阶段，如果之前的策略所产生的收益为（c,y），或（d,w）或（a,z），参与者 1 选择 c；如果之前参与者 1 选择 y 且参与者 1 没有选择 c，参与者 1 会选择 d；如果之前参与者 1 选择 c，参与者 2 没有选择 y，那么参与者 1 会选择 a，否则他会选择 b。阶段 1，参与者 2 选择 y。其他阶段，如果之前策略产生的结果为（c,y）或（d,w）或（a,z），那么参与者 2 会选择 y；如果之前参与者 2 选择 y 且参与者 1 没有选择 c，那么参与者 2 会选择 w；如果之前参与者 1 选择 c 且参与者 2 没有选择 y，那么参与者 2 会选择 z，否则他会选择 x。为得到这一收益总结出满足条件以便得到一个满足子博弈精练纳什均衡。

2. 早在 19 世纪 20 年代，"双缺席"这一惯例就已在议会产生。如果一个议会议员打算在一次正常的选举上缺席，他会预先安排一名对立方的议员与自己保持一致，以便达到两票不参选的目的。为了模式化这一情形，假设有两名议会议员，即议员代表史密斯和约翰（他们因为其他事，将缺席议院的选举）。议员代表史密斯从阶段 1 开始每三个阶段缺席一次（也就是 1,4,7,…阶段缺席），并且缺席议会能得到 3 个单位的收益。议员代表约翰从阶段 2 开始也会每 3 个阶段缺席一次（也就是 2,5,8,…阶段缺席）并且他也能得到 3 个单位的收益。我们把这些议会代表缺席的阶段称为"议会代表的旅行阶段"。每一个这样的阶段都会有议会选举，如果议员史密斯出席议会并参选，他将得到 5 个单位的收益，而如果约翰出席议会并参选，史密斯会损失 5 个单位的收益。同样的，如果约翰出席议会并参选，他将得到 5 个单位的收益，而如果史密斯出席议会并参选，约翰会损失 5 个单位的收益。从而可以得出这一结论：如果史密斯和约翰都出席议会并参选，那么双方都将得到零收益（5−5=0）。在议员的旅行时期，议会代表的收益可以分为以下几种情形（1）如果他未出席议会，其他代表未参选，他所得收益为 3；（2）如果他未出席议会而其他议员参加了选举，此时他的收益为−2（=3−5）；（3）如果双方都出席并参选，收益为 0；（4）如果他出席议会并参选而对方未参选，此时所得收益为 5。每一阶段，议会代表要同时决定是否出席议会并参选。每个议员都在寻找策略以最大化其单一阶段的当前收益，折现系数为 δ。寻找一个子博弈精炼纳什均衡以使每个议员在他的旅行阶段不出席议会并且在这样的阶段他们能做到"双缺席"。

3. 我们现在开始回过头来讨论 14.2 节中关于吸血蝙蝠的博弈，但是现在假设两只吸血蝙蝠成功的概率不同。出于这一目的，现在将两只吸血蝙蝠分别命名为安雅和斯柏克。安雅夜间捕食成功的概率为 a，斯柏克成功的概率为 s。为 14.2 节的一对策略寻找成立的条件以便得到一个子博弈精炼纳什均衡。

4. 三家建设公司 A、B、C 按照惯例竞标政府的铁路建设工程，每一阶段政府拟定一个契约，三家公司同时开始竞标。所谓竞标就是一家公司为了得到这项工程给出的他所能接受的价格，最终价格最低的公司会赢得此项工程。政府规定竞标定价必须以 100 为单位增加（故可行方案为 0,100,200,…）然而，一旦竞标价格超过 2 000，对公司来说这项契约就没有签订的价值。如果两家或者更多的公司提交的竞标方案为最低价格，那么每一家公司都有相同的机会得到此项工程。假设一个公司执行合同要花费 1 000，而这家公司是以 b 的出价竞标成功的，他的收益为 b−1 000。在这场无限次重复博弈中，一家公司的收益就是其收益所创造的现值。

a. 为这个阶段博弈寻找一个对称的纳什均衡。

b. 假设准备竞标的各公司采用"冷酷策略"以交替竞标的方式相互勾结。在阶段 1,4,7,…公司 A 以 2 000 的价格竞标而公司 B 和 C 的竞标价格高于 2 000；在阶段 2,5,8,…公司 B 以 2 000 的价格竞标而公司 A 和 C 的竞标价格高于 2 000；在阶段 3,6,9,…公司 C 以 2 000 饿价格竞标而公司 A 和 B 的竞标价格高于 2 000。假设任何违背此项相互串通协议的行为都将使公司永远的回复到一个对称的阶段博弈的纳什均衡状态。为这个策略所产生的收益指定成立条件以便得到一个子博弈精炼纳什均衡。

5. A 是镇上唯一的商店，它当前所得收益为 20。但是它有一个潜在的竞争对手（B 正在考虑进入市场，开

始竞争)。我们把每一时期分成两个阶段:阶段1是一个进入市场或撤离市场的阶段;阶段2是一个价格竞争的阶段。如果B之前没有进入市场的经历,那么在阶段1即当前阶段他要决定是否进入市场。进入市场首先需要一次性支付成本16。如果B不进入市场,A所得当前收益为20。如果B进入市场,在阶段2两家公司将开始价格竞争,从低、中、高三种定价策略中选择,与之相对应的收益如图14—6所示。

B

A		低	中	高
	低	3,3	5,4	8,2
	中	4,5	7,7	12,4
	高	2,8	4,12	10,10

图14—6

如果B过早地进入了市场,在阶段1他要决定是否退出市场。如果退出市场,他还能找回一半的进入成本。如果不退出市场,两家商店将展开如图14-5所示的博弈。这场博弈是无限次的,而每家公司都试图最大化其当前收益。假设B可以随时进入或者退出市场,对B来说,其退出市场的这一阶段将获得零收益。

a. 考虑如下策略组合:对A而言,如果B参与市场竞争,并且这是B在市场上的第一个阶段或者A在之前B参与市场竞争时定了较低的价格,那么此时所定的价格较低,否则价格处于中等水平。对B而言,考虑阶段1首先该做什么,第一阶段不进入市场,其他时期有以下几种情况:(1)如果之前没有进入市场的经历则不进入市场;(2)如果之前有过参与市场竞争的经历,一旦A将价格定得很低(此时B也在市场上)B退出市场,否则不退出。阶段2,如果B在阶段1进入了市场,那么其所定价格处于中等水平。我们注意到这一策略阻止B进入市场,因为B预想到一旦B进入市场,A会采取侵略性的定价策略。找出满足这一策略的条件以产生一个子博弈精炼纳什均衡。

b. 寻找一对让B进入市场的策略,并且找出满足这一策略的条件以产生一个子博弈精炼纳什均衡。

6. 回过头来考虑一下14.3中借款给国王的博弈,假设在阶段1,国王需要的借款为M,而未来需要的借款仅为100。我们设想国王和放款人都使用了如图14—4所示的策略。

a. 找出满足这一策略的条件以产生一个子博弈精炼纳什均衡。

b. 证明要达到均衡需要满足如下条件,最初的借款不能高于任何预期的将来借款。解释一下为什么这种论断是正确的。

7. 正如伊丽莎白·雅顿(Elizabeth Arden,伊丽莎白·雅顿品牌创始人)所说:"名誉源于重复,而名誉吸引了顾客。"让我们一起见证一下这条名言是如何发挥作用的。每一阶段,生产者面临两种选择:生产低质产品以10元每单位的价格出售,还是生产优质产品以20元每单位的价格出售。假设购买一件劣质商品,消费者所能承受的最高价格为15,而为了购买一件优质商品,消费者愿意承受高达50的售价。问题在于,消费者不能透过商品的外表区分质量的好坏。然而,一旦购买该商品消费者可以通过使用辨别其质量的好坏。我们认为每一阶段都有一个消费者,并且只要商品是有剩余价值的,消费者就会购买。所谓剩余价值就是消费者所期望得到的商品价值(如果消费者期望商品是劣质的,他所能承受的价格为15,而如果期望是优质的,则为50)低于所支付的价格。生产者的收益就是其所得利润的现值。任何一个阶段,没有消费者购买商品生产者所得收益为零,如果消费者购买商品,商品售价要降到成本以下。生产者的折现系数为δ。每一阶段,生产者都要决定生产何种质量的商品以何种价格出售。消费者通过比较商品的价格(而不是质量)决定是否购买此种商品。假设每一阶段所有过去的选择包括价格、质量、最终购买等都是共同知识。寻找一个策略使得生产者总能生产高品质的商品并能以50的价格出售给消费者。找到满足这一策略的条件以产生一个子博弈精炼纳什均衡。

8. 我们修改一下上个问题,假设即使购买了该商品也不能完全正确评价其质量。每一阶段,生产者选择商品的耐用性,生产的商品或者能用或者不能用,但是耐用性决定了商品是可以用的。生产者可以以20的成本生产一件持久耐用的商品并且这一商品能正常使用的概率为90%,也可以以15的成本生产一件不太耐用的商品,此种商品能正常使用的概率仅为60%。消费者愿意花费30购买一件能正常使用的商品却不愿花一分钱购买一件不能用的商品。如前所述,假设每个阶段有一个消费者,他决定是否购买商品,并且他在购买的时候不能观察到商品的耐用性,买回去后也只能知道它是否能正常使用。

如果不买，不需要支付任何费用。如果购买，他的收益等于期望的商品的价值减去他支付的价格。假如商品有很高的耐用性，期望价值为 0.9×30+0.1×0=27，而如果商品的耐用性很低，期望价值为 0.6×30+0.4×0=18。每一阶段，生产者要决定生产何种耐用性的商品以何种价格出售以便最大化他的期望收益。当前阶段，如果消费者不购买商品所得收益为零，如果购买所支付的价格低于生产成本。考虑一下如下策略：对生产者而言，在阶段 1，生产一个耐用性很高的商品并以 27 的价格出售；其他阶段，如果之前生产的商品都能正常使用那么接着生产耐用性很高的商品并以 27 的价格出售；否则生产不太耐用的商品以 18 的价格出售。对消费者而言，阶段 1，售价不高于 27 时购买该商品；其他阶段：（1）如果之前购买的商品都能正常使用，只要售价不高于 27 就购买；（2）如果之前购买的商品有些不能正常使用，售价不高于 18 时购买。找到满足这一策略的条件以产生一个子博弈精炼纳什均衡。

9. 我们回过头来分析 13.3 小节中所述的无限次重复堑壕战博弈。考虑如下策略：在阶段 t 时，假设 a^t 和 g^t 分别代表协约国一方及德军方面的行为，我们把 C_A^t 定义为早于阶段 t 的时期，此时（a^t，g^t）=（击中，不击中），也就是说，协约国一方蓄意射击而德军方面致力于维持和平。类似的，将 C_G^t 定义为早于阶段 t 的时期，此时，（a^t，g^t）=（不击中，击中）。接下来，我们定义 $C^t=C_A^t-C_G^t$，C^t 就是一个计数器，记录着协约国一方比德军方面多违背了几次休战协定。所谓违背休战协定，意为一方击中而一方却不击中。令 $C^1=0$（也就是说从 0 开始计数），考虑以下对称的策略，要求博弈双方的行为在计数范围内。首先定义协约国一方的收益，阶段 1 选择不击中对方；在阶段 t（$\geqslant 2$），如果 $C^\tau \in \{-1,0,1\}$ 且 $\tau \leqslant t-1$，同时 $C^{t-1}=0$，或 $C^{t-1}=1$，选择不击中对方。在阶段 t（$\geqslant 2$）如果满足以下条件：（1）$C^{t-1}=-1$ 且 $C^\tau \in \{-1,0,1\}$ 其中 $\tau \leqslant t-1$ 或者（2）$C^\tau \notin \{-1,0,1\}$，有时满足条件 $\tau \leqslant t$，那么选择击中对方。接下来定义德军方面的收益，阶段 1 选择不击中对方，阶段 t（$\geqslant 2$）如果 $C^\tau \in \{-1,0,1\}$ 且 $\tau \leqslant t-1$，同时 $C^{t-1}=0$ 或 $C^{t-1}=1$，选择不击中对方。阶段 t（$\geqslant 2$）如果满足以下条件（1）$C^{t-1}=1$ 且 $C^\tau \in \{-1,0,1\}$ 且 $\tau \leqslant t-1$ 或者（2）$C^\tau \in \{-1,0,1\}$ 有时满足条件 $\tau \leqslant t$ 那么选择击中对方。这一策略让双方参与者开始选择维持合作关系。假设所记录的数据一直都是−1,0 或者 1，所以一直以来双方背离休战协议的次数是大体相同的。如果 $C^{t-1}=0$，这意味着过去双方背离休战协议的次数几乎是相同的，双方都会保持合作关系选择不击中对方。如果 $C^{t-1}=1$，这意味着协约国一方比德军方面违背协议的次数多，此时协约国一方士兵选择不击中对方而德军士兵选择击中对方，为了获得胜利，德军方面甚至突然袭击协约国一方，我们注意到结果是 $C^{t+1}=0$。如果 $C^{t-1}=-1$，这意味着德军方面比协约国一方违背协议的次数要多，此时协约国士兵选择击中对方而德军方面选择不击中对方，这样 $C^{t+1}=0$ 成立。最后，如果所记录的次数不是−1,0,1，那么违背协议的次数多 2 甚至更多，此时在所有给定的时期，双方都会选择击中对方。换句话说，如果过去，双方违背协议的次数有很大差距，那么双方会永远放弃休战。找到满足这一策略的条件以产生一个子博弈精炼纳什均衡。

10. 每一阶段，两个参与者要同时在上和下之间做出选择。如果双方都选择上，那么每个参与者都有 80% 的可能获得 10 个单位的收益，20% 的可能获得 3 个单位的收益。如果双方都选择下，每个参与者一定能获得 5 个单位的收益。如果一个参与者选择下，另一个选择上，那么选择下的参与者一定能获得 10 个单位的收益而另一方有 60% 的可能获得 10 个单位的收益，40% 的可能获得 3 个单位的收益。每当做出选择的时候，每个参与者都知道自己过去的选择及双方的收益，但是谁都不能明确地知道对方过去的选择。这一阶段博弈是无限次重复的。每个参与者的折现系数都是 δ。考虑以下对称的策略：阶段 1，两个参与者都选择上；其他阶段，如果双方过去所得收益相同（3,5 或者 10），那么选择上，否则选择下。找到满足这一策略的条件以产生一个子博弈精炼纳什均衡。

第 15 章

非固定居所机构的互动博弈

学习目标

15.1 世代交替间的合作博弈

15.2 多人合作博弈

本章导读

在这章中，合作将会在最不可能的条件下实现。从 13 章和 14 章我们知道，合作博弈是建立在参与者对未来前景展望的基础上。未来预期使得有些人产生不作为行为，而这种结果产生的威胁可以导致其现在的积极作为。**只要参与者未来还有博弈的可能，还在乎他们的未来收益，合作就可以一直持续下去。**

但很明显，也存在另一种情况，参与者好像看不到“未来的影子”。比如，思考一下一名雇员将要退休的例子。因为他并不期望在未来和他现在的同事有往来，那么在他们之间还存在合作的基础吗？或者思考一下一个人被要求照顾她年迈的父母这个例子。如果她不履行家庭义务，她的父母怎么可能去惩罚她？当然，儿女可能会出于爱或者害怕失去继承资格而尽赡养义务。但如果这些外在因素不存在，那是否就意味着当父母老了，无法继续付出的时候就会被遗弃呢？

在 15.1 小节，我们思考在人们意识到他们和其他人的合作在某一刻将终结的场景。在模拟这些有限的生命时，一个重要的假设是参与者的生命阶段不能完美地重合。在任何阶段，有的人处于青少年阶段，有的人处于中年阶段，有的人已到暮年。有些人会离开（去世了，或者离开组织）同时又会有新成员加入。尽管有人离开了，群体会得到不断的补充，在这种情况下，组织里的人永远处于互动状态。我们就会明白，只要组织无限存在，具有有限生命的参与者的合作就会一直持续下去。合作可能不会持续整个生命过程，但这已使故事变得有趣。

对现阶段的合作来说，另一个具有争议的经典例子是两个陌生人相遇时的情况，因为在这两个人中，其中的一个不可能使这个场景重复发生。当两个相遇的参与者并不期望在未来还会再见面时，我们称这是个“大的”群体。所以，当杰西卡不合作时，如果詹姆斯威胁惩罚杰西卡，合作就不会保持，因为杰西卡并不认为他们在未来还会再见面。而且，尽管一个人并不能规范另一个人的行为，整个社会却可以实现这个功能。例如，

PDG

如果未来和杰西卡有往来的人知晓杰西卡和詹姆斯过去的情况，他们就可以通过对詹姆斯不友好来惩罚杰西卡。

以下讨论假设相遇的信息是公开的，有时也并非这种情况。然而，我们会看到在实际生活中信息如何成为日后社会惩罚那些表现不好的人的基础。这都会在 15.2 小节探讨，在 15.2 节中，我们证明中世纪的商人信任彼此，尽管他们日后可能没有经济往来，这是因为害怕不良行为会被报告到法官那里。尽管法官并没有执法能力，他的裁决，也就是对于不良商业行为的判断，却对未来商业伙伴有持续的影响力。这个案例同样适用于 eBay 网。

15.1 世代交替间的合作博弈

孩子是我们可以确定的唯一不朽的形式。

——彼得·乌斯季诺夫（*Peter Vstionov*），

奥斯卡最佳男配角

一些人老去，一些人降生，但社会却一直持续下去。公司里老员工退休，但同时，公司又招来新员工代替老员工。公司可以持续存在发展尽管公司的创始人不能始终相随，立法成员因为任职到期或者再次选举落选面临职业生涯终结，但她的职位会由其他人取代。一个政治家去世了，立法组织却还依旧存在。实际上，这类案例通常用来描述个人组成的集体，不管是大家庭、大学生联谊会、球队或是部落。

在这一小节中，我们探讨组织中的合作。在这个组织中，所有成员的生命都是有限的，而组织形式一直持续，因为新招录的成员会不断补充组织。这种补充并不是一下全部完成，而是部分地随着时间持续完成。在每个阶段，一些人离开了，被取代了，而组织中的其他人还处在即将离开组织的时刻。换句话说，它就像一个大家庭，在这个家有孩子、父母、祖父母，可能还有曾祖父母。孩子长大之后成为父母，父母成为祖父母，祖父母，甚至成为曾祖父母，直到去世。

不同代际的人共存的群体结构被称为世代交替。因为我们即将利用世代交替模型来做广泛的研究，对群体结构有一个清晰的把握是有必要的，这儿举一个有用的例子来说明。观察图 15—1，我们发现在阶段 9 有三个人一起生活：雅各布，艾米莉和迈克尔。雅各布处于生命的第 3 也是最后一个阶段，艾米莉正经历中年危机，迈克尔还是个新生命。到阶段 10，雅各布已过世，艾米莉现在成为这一组织的老年人，而迈克尔正值中年危机，艾玛取代了雅各布的位置。在每一个阶段，有人去世，有人出生，所以组织总是有三个人，每代一个人。

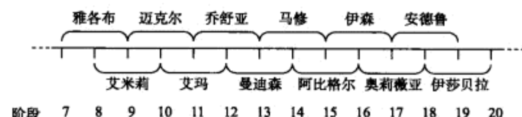


图 15—1 世代交替的一个团体

从前面章节中我们知道，当参与者能预计到互动的次数有限（不论是 2 次，10 次还是 20 次）合作就很难达成。这层逻辑的一个隐含条件是不会有人在他们生命的最后阶段参与合作（这里的“生命”是指在组织中的“生命”）。然而，事实是一个生命力无限的世代交替组织可以提供许多合作的机会。这一小节将证明合作如何在一个生命力无限的社会中持续即使社会中的成员都是由有限的生命组成。

第一个应用和部落或村庄有关，在部落或村庄，每个人被要求为公益事业做贡献。对个人来说，成本很大，但如果每个人都这样做的话，部落的所有成员的情况都会好转。唯一的挑战就是避免有些人逃避责任，利用他人的奉献成果。第二个例子是普通的代际环境，在这个环境下，一个老人希望他成年的孩子将来能够照顾他。假设他们之间没有爱，孩子会不会给年迈的父母经济援助？第三个也是最后一个例子来源于政治界，它涉及的是一位无能的当选官员在削弱所在党继续执政机会的同时，追求自己的利益。

◯案例研究

部落防御

任何组织的一个共同特点就是它所有的成员都需要为集体的利益做贡献。一位为组织利益努力工作的成员使全体成员获益，而努力工作的代价仅仅由这一个成员承担。如果每个人都做出贡献，那么由每位成员实现的

利益就超过个人的付出。所以，如果都努力工作，每个成员得到的利益要大于每个人都偷懒得到的利益。唯一的挑战是就避免“搭便车现象”——有些成员逃避自己的责任，又享受着其他成员努力奋斗的果实。当然，如果搭便车的人足够多，那么所有的成员都放弃为集体做贡献，这样集体的利益就会损失很大。

这种场景下的一个例子就是防御机制。不管是一个国家、一个部落或一个村庄，每个成员都努力抵御敌人入侵，这使组织的所有成员获益。但代价（不论是增强了人的身体防御能力还是在战争中付出自己的生命）都是为了集体的利益付出的个人成本。让我们探索一种策略可以使部落成员都做出贡献从而避免那些“逃避责任”的现象发生。

为了把复杂的问题简单化，假设一个部落有 $N \geqslant 2$ 个成员，每个成员都只有两个选择，奉献或者逃避责任。每个人奉献的成本是 10，而逃避责任的成本为 0。N 个成员的每一个从 m 个成员的奉献中得到的利益为 $6 \times m$，所以每增加一个奉献的成员都可以使每个人的收益值增加 6。用 V（奉献，m）和 V（逃避，m）分别代表社会中成员选择奉献或者逃避得到的收益值，假定 m 个成员都选择做贡献：

$$V(\text{行动}, m)=\begin{cases}6(m+1)-10 & \text{行动}=\text{奉献}\\ 6m & \text{行动}=\text{逃避}\end{cases}$$

首先注意逃避是一种主流选择。每个成员从自己(和其他成员）的奉献中得到的利益为 6，但他付出了的成本为 10，所以他的收益降低了 4。他的奉献使组织的收益增长了 $6N$，大于个人付出的成本 10（因为 $N \geqslant 2$）。那么从个人的角度出发，每个成员都想要逃避责任，但如果每个人都这么做，那么每个人得到的收益为 0。从组织的角度出发，每个人都做出贡献是最好的选择，因为这样每个人得到的收益为 $6N-10$。

如果成员之间的互动是无限的，从前面一章我们知道一个简单的解决办法。思考一种策略，在这种策略下，只要所有的成员在过去都一直在做贡献，每个成员都做出贡献并且一直这样贡献下去。如果有一个逃避责任，惩罚就是每个人都逃避一段时间，直到他们都返回合作状态中。只要：

$$(6N-10)+\delta\times(6N-10)+\delta^2\times(6N-10)+\cdots$$
$$\geqslant 6(N-1)+\delta\times 0+\delta^2\times(6N-10)+\cdots$$
$$\Rightarrow (6N-10)+\delta\times(6N-10)\geqslant 6(N-1)$$
$$\Rightarrow \delta \geqslant \frac{4}{6N-10}$$

这个解决方法太简单，所以我们假设每个成员都认为自己的生命是有限的，当组织有 $T \geqslant 2$ 个阶段时，情况尤其如此。在生命中的任何一个阶段，每代都有 N/T 个成员，假设 N 是 T 的倍数，N/T 是个整数。比如，如果 N=100，T=4，那么，在任意一个阶段，有四代人，每代人有 25 个成员：25 个人处于生命中的第一个阶段（“孩子”），25 个人处于第二阶段（“青年”），25 个人处于第三阶段（“成年”），25 个人处于生命的最后阶段（“老人”）。每个成员的收益就等于在生命中的 T 个阶段中各个阶段收益的简单相加（换句话说，δ=1）。

想想看，努力保持每个人都做贡献的合作状态。一个急需解决的问题就是没有办法使老年人做出贡献。他们已经处于生命的最后阶段，这就意味着如果他们逃避责任也不会得到惩罚，因为他们已经没有未来。那么我们就只能妥协，允许处在生命最后阶段的人逃避责任。

但是我们可以使其他成员做出贡献吗？思考一种策略使得每个人都在生命的前 T-1 的阶段付出，在最后一个阶段逃避责任。如果过程中有偏差，特别是如果有些不处于最后阶段的人不付出，那么每个人都停止付出一段时间再回到做出贡献的状态（当然，除了那些处在生命中最后阶段的成员）。

在证明这种情况是一个均衡的过程中，首先注意对处在生命最后阶段的人来说，这个策略是最优策略，因为他就是打算逃避责任。接下来，考虑处在倒数第二阶段的成员，她的策略是做出贡献，当：

$$\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta\times 6\left(\frac{T-1}{T}\right)N \geqslant 6\left[\left(\frac{T-1}{T}\right)N-1\right]+\delta\times 0 \Rightarrow \delta \geqslant \frac{2}{3N\left(\frac{T-1}{T}\right)N} \tag{15.1}$$

时，她的策略是最优策略。这个成员期望其他任何比在生命最后阶段年轻的人都做出贡献，在一个组织中，包括她自己，做贡献的人有 $\left(\frac{T-1}{T}\right)N$ 个。通过奉献，这个成员就可以得到的收益为 $6\left(\frac{T-1}{T}\right)N-10$。在她生命的最后阶段，她获得的收益为 $6\left(\frac{T-1}{T}\right)N$，像先前那样，只要那些不是处在生命最后一个阶段的成员都在付

出，而她，现在处在生命的最后一个阶段，则可轻松度过。这两个收益之和就是不等式（15.1）的左边。再看不等式的右边，如果这个成员在现阶段选择不付出，她当前的支付收益就提高了 4，从 $6\left(\frac{T-1}{T}\right)N-10$ 提高到 $6\left[\left(\frac{T-1}{T}\right)N-1\right]$；总体利益下降 6（因为她不付出），但她也避免了付出成本 10。逃避的代价就是她的行为可能导致下一个阶段的人也不付出，这样她得到的收益就为 0。如果不等式（15.1）成立，对处在 $T-1$ 阶段的成员来说，通过付出得到的收益就大于逃避责任得到的收益。

接下来，思考那些处在 $T-2$ 阶段的成员的情况。均衡条件就变为：

$$\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta\times\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta^2\times 6\left(\frac{T-1}{T}\right)N \geqslant 6\left[\left(\frac{T-1}{T}\right)N-1\right]+\delta\times 0+\delta^2\times 6\left(\frac{T-1}{T}\right)N \Rightarrow \delta\geqslant\frac{2}{3\left(\frac{T-1}{T}\right)N-5} \qquad (15.2)$$

逃避意味着现阶段获得一个更高的收益 4，但这个成员在下一个阶段会有损失，因为每个人都选择逃避。这就意味着在下一个阶段损失的收益为 $6\left(\frac{T-1}{T}\right)N-10$。

对任何比 $T-2$ 阶段年轻的成员来说，条件式和（15.2）相同：逃避使现在的收益提高了 4，但是在下个阶段会损失 $6\left(\frac{T-1}{T}\right)N-10$ 个收益，因为所有的参与者在这个阶段都停止付出。在接下来的阶段中，得到的收益就和付出时一样，因为所有的参与者又开始付出。所以，当（15.2）条件式满足时，所有比 $T-1$ 阶段年轻的成员发现最优策略就是付出。

总之，如果不等式（15.1）和不等式（15.2）都成立，这种策略组合就是子博弈精炼纳什均衡。注意如果满足不等式（15.2）的条件，那么也满足不等式（15.1）的条件。从直觉来讲，对那些处在倒数第二阶段的成员来说，欺骗的诱惑力很低是因为欺骗将导致她损失 $6\left(\frac{T-1}{T}\right)N$ 个收益。每个成员在最后一个阶段都会得到很好的照顾（她开始轻松的生活而其他更年轻的成员在努力付出），但如果她在倒数第二个阶段偷懒，就不会得到“退休福利”。真正的困难是引导那些离退休还早的年轻人去付出、去贡献。

15.1 理解测验

在前面的策略组合中，部落防御的惩罚机制是所有的成员在一个阶段都不付出。现在，假设惩罚的时间是偷懒的那个成员的所有阶段。所以，如果一位成员在她生命中的 t 阶段逃避（当她应该付出时），那么所有的参与者在接下来的 $T-t$ 个阶段都不付出。推导能保证所有人在他们生命中的前 $T-1$ 个阶段都做贡献的条件式。


点拨

即使参与者的生命是有限的，如果群体不断被新生代补充，合作至少可以部分维持。一个人可以获得奖励（因为合作）和被下一代惩罚（因为背叛）。

照顾好你年迈的父母

当我们还是孩子时，我们常常想当我们长大了，我们就再也不脆弱。但是长大就意味着接受脆弱…… 活着就意味着脆弱。

——马德琳·英格（Madeleine L'engle），《水上行走》（*Walking on Water*）

时刻善待你的孩子，因为他们是为你选择疗养院的人。

——菲利斯·迪勒（Phyllis Diller），美国著名喜剧女演员

如果孩子爱父母，或者有一大笔遗产做赌注，这很容易理解他们为什么会照顾年迈的父母。但是假如这两个条件都不具备呢？假设一个人是无情的，生下来就没有照顾她贫寒年迈父母的心，这个人可以被引导做正确的事，照顾她年迈的父母吗？

为了探讨这个答案，假设一个人经历两个成年阶段：成年和老年。为使这场讨论更容易被理解，我们假设成年人为一位女性，而老年人为男性。不论年龄，一个人在任何阶段的收益都取决于他或她的消费额。消费额和收益的关系如图 15—2 所示：

PDG

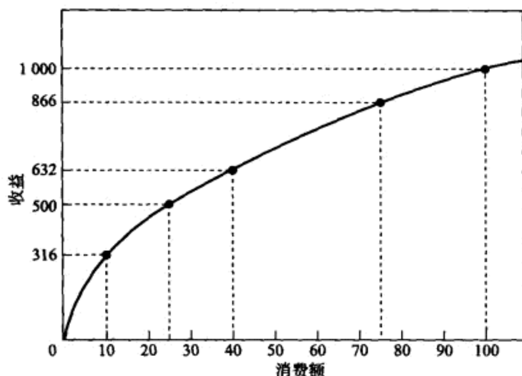


图 15—2 消费额与收益的关系

请注意，消费得更多（买更多的衣服，在一个不错的餐馆吃更多的肉，从 iTune 上下载更多的资源）意味着更高的收益。然而，当一开始的消费更高时，从更高消费额中增加的收益就越少。比如，当消费额从 10 增加到 25 时，增加的收益为 184，而消费额从 25 增加到 40 时（同样是 15 个增长点），收益增加只有 132。换句话说，你第一次从 iTune 上下载的歌比你的第 100 次下载更有价值，而第 100 次下载又比第 1 000 次下载更有价值。

作为一个成年人，她的创造体现在两个方面：工作收入 100，生一个孩子。所以，在每个阶段有三代人：年轻人，成年人和老人。现阶段的年轻人会成为下一个阶段的成年人，成年人会成为老人，而令人悲伤的是，老人将去世。每个人生活三个阶段。伤脑筋的是老人不赚工资，为了使问题简单化，假设他们没有任何积蓄，老人就依靠他的孩子（一个现在赚钱的成年人）来养活他。

在任何时候决策者都是这个成年人，她必须决定她收入的多少用来消费，多少用来养父亲。为了简单，我们忽略年轻人的花销。

代际准则是如果父母遵守规则，成年人就应该照顾父母。“照顾好父母”就意味着至少提供 25 个消费额。“遵守规则”就意味着这个人必须：（1）照顾好她的父亲；或者（2）她的父亲不遵守规则。例如，当她父亲照顾她的祖父母，如果她不照顾她的父亲就是不遵守规则。如果她父亲没有照顾她的祖父母（而她的祖父母遵守规则应该得到照顾），那么如果她没有照顾她父亲就不算违反规则。合理性在于惩罚她的父亲不照看她的祖父母。代际准则的本质就在于一个人有义务照看父母，除非她的父母忽视她的祖父母不管不问，得到的惩罚就是也被忽略不管。

这个准则反映了家庭成员间的期望。然而，我们仍然需要规定个人策略。假设策略规定：如果你的父母遵守规则，你就必须提供 25 个消费额给父母；如果你的父母不遵守规则，就不必提供消费额。注意如果每个人都采取这个策略，假设第一个人是遵守规则的，接下来所有的后代都会遵守规则。所以，每个成年孩子都会照顾她年迈的父母。

接下来的一步是推导出这个对称策略组合成为子博弈精炼纳什均衡的条件式。假设一个人没有必要照顾她的父母。可能她的父母并没有养活她的祖父母，即使她的祖父母养活了她的曾祖父母。因为这个成年人的孩子提供 25 个消费额给她，不管她提供了多少给她的父亲，她应该会转移策略使她在当前阶段的收益最大化，那就是不管她年迈的父亲。从图 15—2 我们知道，她的收益是 $1\,000+\delta\times500$，因为她从当前消费额 100 得到 1 000 个收益并且在下个阶段通过消费额 25（来源于她的孩子）获得 500 个收益，这里她的折现系数为 δ。所以，这个策略明显是最优策略。换句话说，如果规则说你可以不赡养你的父母，那么不赡养就是最优策略。

更复杂的问题是当一个成年人必须赡养她的父亲时。比如，假如她的父亲供养了她的曾祖父。按照策略的规定，她应该提供 25 个消费额。她没有理由提供比 25 更高的消费额，因为这样做只会降低收益而在未来不会产生更高的收益。所以，25 个转移明显优于大于 25 个转移。因为这意味着当前有 75 个消费额（866 个收益）未来有 25 个消费额（当她老的时候她的孩子提供给她 500 个收益，如图 15—2）。现在考虑她提供给她父亲的消费额低于 25。因为这会导致她的孩子以后不赡养她（不管她的转移是 0 还是 24）她宁愿选择 0 而不是 0~25 之间的任意值。不赡养她父亲得到的收益为 1 000，虽然她当前消费 100，但她在下一个阶段什么都得不到，她的孩子会因为她不赡养老人惩罚她。那么当：

$$866+\delta\times500\geqslant1\,000\Rightarrow\delta\geqslant\frac{134}{500}=0.268。$$

时，成年人发现赡养父母是最优策略。

这个观点再简单不过：当一个人的父母老了，也没

有遗产作为诱惑的时候，他的孩子不赡养他，他也不能惩罚他的孩子。而这里的惩罚机制的决定权不在年老的父母，而在他的孙子。如果一个成年人不赡养她的父母，那么她的孩子会以此为理由不赡养她。每代人都会赡养父母履行社会责任，因为他的下一代以此为依据赡养他。所以，年迈的父母都会被自私的孩子赡养。

执政党和"跛鸭"总统

政治家就像尿布，要勤洗勤换才好。

——佚名

执政党为了当选可能会持温和观点。但如果当选的是一个无能的人，会发生什么情况呢？如果他缺乏连续当选的素质，他会一门心思采取激进政策损害他所在党的名誉甚至保留执政位置的前景吗？这里我们证明如果这个无能的总统在他离职后至少关注一个阶段的政策，那么对所在党名誉的关注能使他避免采取激进政策。

假设有 5 个不同的平台或政策：自由主义，适度宽松主义，适度主义，适度保守主义和保守主义。民主党的意识形态是自由主义，而共和党的是保守主义。然而，在大选中，最优策略是适度政策。在表 15—1 可以看出，它列举了在不同策略下，民主党赢得大选的概率（1 减去共和党赢得大选的概率）。一个政党采取的策略离选民理想的适度策略越远，这个政党赢得大选的概率就越低。比如，如果共和党采取保守立场，那么民主党如果采取自由主义策略（自由主义策略和保守策略与适度政策的距离一样远）就有 50% 的机会获胜，而如果民主党采取与选民的适度政策更相近的适度宽松政策，那么赢得大选的概率为 70%。

表15—1 赢得大选的概率

民主党的平台	共和党的平台	民主党赢得大选的概率
自由主义	适度政策	0.1
自由主义	适度保守主义	0.3
自由主义	保守政策	0.5
适度宽松政策	适度政策	0.4
适度宽松政策	适度保守主义	0.5
适度宽松政策	保守政策	0.7
适度政策	适度政策	0.5
适度政策	适度保守主义	0.6
适度政策	保守政策	0.9

尽管假设候选人关注任职期的政策是很自然的事情，但为使事情简单化，我们假设他们只关注现在实施的政策。表 15—2 列出了他们的政策偏好。实施的政策越保守，民主党得到的收益就越低。例如，民主党从适度宽松政策得到 4 个收益，而从适度保守政策得到的收益为−3。

表 15—2 政治家的收益

政策	民主党的收益	共和党的收益
自由主义	5	−5
适度宽松政策	4	−3
适度政策	0	0
适度保守主义	−3	4
保守主义	−5	5

假设任职期限规定当选政府只有一届任期。并且，一个政治家只生活两个阶段，更准确来说，她只关注这两个阶段实施什么政策。你可以想象她在过了两个阶段之后变得不关心政治的情景，看看美国有线电视，或者看看《美国偶像》。从政治家的角度来看，她在第一阶段得到的收益取决于她实施的政策，在第二个阶段得到的收益取决于她的继任者实施的政策。后者用 δ 来表示，$0 \leqslant \delta \leqslant 1$，这一条件取决于平时对偏爱政策的讨论或者是当她不在任职期对政策关注少的讨论。

我们分析的第一步就是证明一个政党想要通过实施比政党意识形态更温和的策略来赢得名誉。例如，假定选民期望共和党实施适度保守政策。如果选民期望民主党实施自由主义政策，那么，在大选期间，民主党的期望收益就是 $0.3\times5+0.7\times(-3)=-0.6$。由于采取比对手更激进的政策，民主党只有 30% 的机会获胜，尽管在那些情况中，他们会因为实施自由政策获得 5 个收益。共和党有 70% 的机会获胜，实施适度保守政策从而使民主党的收益为−3。相反，如果民主党可以说服选民，保证他们不实施极度自由主义政策（他们将会实施适度自由政策），那么他们的预期收益就会相对高一点，为 $0.5\times4+0.5\times(-3)=0.5$。即使选民相信共和党会实施保守策略，对民主党来说，想提升党的名誉，实施适度宽松政策比自由主义政策要好，因为：

$$0.7\times4+0.3\times(-5)>0.5\times5+0.5\times(-5)\Rightarrow 1.3>0$$

对共和党来说，这个分析同样适用；他们也宁愿选择采取比他们党意识形态更温和的策略来提升党的名誉。

如果不关心是否能赢得大选，一个政党会实施他们党的意识形态政策，所以民主党会实施自由主义策略，

共和党会实施保守主义策略。如刚才所示，选举压力使政党把自身变得在选民眼里更温和，因为这样做会提高赢得大选的机会。接下来要做的就是发展一种机制，这种机制对一个党在表示自己可以实施更温和策略时更有帮助。

在构建两党实施比他们意识形态更温和的策略的均衡中，思考下面这个策略：一开始民主党的名誉就是实施适度宽松政策，接下来他还会实施适度宽松政策。只要前几届的民主政治家实施的是适度宽松政策，当选的民主政治家也会实施适度宽松政策（因为选民期望他这样做）。然而，如果任意一个民主政治家不实施适度宽松政策，那么选民不会再相信民主党的温和策略，民主党的策略就会变为实施自由主义策略。同理可知，共和党一开始的政策是实施适度保守策略，只要他们在过去一直这样做，他们还会继续实施适度保守策略（所以选民继续相信共和党不会实施激进政策）。只要一位当选的共和党人不实施适度保守策略，共和党就会实施保守策略，而这个转变是选民能预料到的。

如前面所描述的，这个党的策略和触发策略很相似。策略实施的困难就在于现在在职的执政党成员只关心她所在党在这两个阶段的收益。在离职后的一段时间她或许还关心实施的政策，但一个阶段过后，她就不会关心现在的政策，也不在乎她所在党的名誉。

为了证明这一对政策是一个子博弈精炼纳什均衡，我们需要演示此均衡规定的最优策略。为了研究，所有的历史都可以被划分为4个部分。（1）两个党派中都没有损害所在党的名誉的情况（就是说，民主党总是实施适度宽松政策，共和党总是实施适度保守政策）；（2）一位共和党人士偏离适度保守政策，而民主党一直实施适度自由主义政策；（3）民主党曾经偏离适度自由主义政策，而共和党一直实施适度保守政策；（4）两个党派曾经都损害了所在党的名誉。因为博弈和策略组合都具有对称性，所以只需求出任何一个党派的最优策略组合即可，下面我们以民主党为例。

在任何阶段，唯一的决策者是当选者，所以假设当选者是民主党派人士。我们将从上述情况（1）进行讨论，所以现在两个党派都被期望实施比他们的意识形态更温和的策略。策略规定了必须采取适度自由主义政策，对当选的党产生的收益为：

$$4+\delta\times[0.5\times4+0.5\times(-3)]=4+0.5\delta \qquad (15.3)$$

她现阶段的收益为4。在下一个阶段，选民期望民主党可以实施适度宽松政策（由于当选者的行为还没有使选民失望），期望共和党可以实施适度保守政策。考虑到那些选民的期望，每个党派有50%的机会赢得大选，在这种情况下，下个阶段的政策为适度宽松政策的可能性为0.5（当前执政党的收益为4）；下个阶段的政策为适度保守政策的可能性为0.5（当前执政党的收益为−3）。

为了确定实施适度宽松政策是不是最优策略，那就考虑一下实施其他政策。如果实施其他政策，那么根据民主党的策略，这个党派将会实施自由政策。合理性在于选民期望他实施他们政党的意识形态政策，因为如果实施的政策与选民的期望有出入就会损害他们政党的名誉。因为未来的收益与现行政策没有关联，当选者就应该选择一种使她现行收益最大的政策，这意味着实施自由主义政策。那样的话她的收益为：

$$5+\delta\times[0.3\times5+0.7\times(-3)]=5-0.6\delta \qquad (15.4)$$

她在任职期间得到的收益为5，在下一个阶段的预期收益为 $0.3\times5+0.7\times(-3)$。因为现在选民期望民主党实施自由主义政策，共和党实施适度保守政策，民主党就只有30%的机会赢得大选。

假设两个党派以前都没有损害过党派的名誉，当表达式（15.3）和表达式（15.4）满足下列条件时，对民主党来说，最好的策略就是实施适度宽松政策。

$$4+0.5\delta\geqslant5-0.6\delta\Rightarrow\delta\geqslant\frac{1}{1.1}\approx0.91 \qquad (15.5)$$

接下来，我们分析情况（2），根据一位共和党人士曾经偏离适度保守政策，而民主党一直实施适度自由主义政策分析。因为共和党人士已经因为偏离适度政策损害他们党的名誉，选民的期望就成为共和党实施保守政策，民主党实施适度宽松政策。当：

$$4+\delta\times[0.7\times4+0.3\times(-5)]\geqslant5+\delta\times[0.5\times5+0.5\times(-5)]$$
$$\Rightarrow\delta\geqslant\frac{1}{1.3}\approx0.77 \qquad (15.6)$$

时，一位在职的民主党人发现实施适度自由主义政策是最优策略，正如她所在党的策略规定的那样。这样做使现在的收益为4，而且满足了选民的期望，这意味着民主党在下一个阶段赢得大选的机会为70%。相反，通过选择实施自由主义政策（是最好的替换政策），她会损害民主党的名誉。在接下来的阶段，选民将期望民主党实施自由主义政策，共和党实施保守主义政策，所以每个党派有50%的机会赢得大选。

为了完成这个分析过程，我们需要考虑情况（3）和（4），这两种情况的共同点是过去民主党都偏离了适度自由主义政策损害了民主党的名誉。选民期望他们实施自由主义政策，这和民主政治家做了什么毫无关系。所以，在职的民主党意识到，不管她在职时实施什么政策，她在下一个阶段的预期收益还是不变，在情况（3）中是 $0.3\times5+0.7\times(-3)$，在（4）中是 $0.5\times5+0.5\times(-5)$。对她来说，使她当前收益最大化的政策就是最优策略，当然，这个政策是自由主义政策。因为这正是民主党策略所规定的，对情况（3）和（4）来说，就是最优策略。

总的来说，我们已经证明如果表达式（15.5）和表达式（15.6）成立，民主党的策略就是最优策略。同理可证，我们可以证明在相同的条件下共和党的策略是最优策略。只要一个政治家足够在乎往她任期后得到的收益，对这两个政党来说，实施比他们意识形态更温和的政策就是一个均衡。通过实施相对适度政策获得名誉，一个政党赢得大选的机会就会增加。但记住这项政策实际上不是由政党决定而是由当选者决定是很重要的，而这个当选者和这个政党不同，她的生命是有限的。不过，生命有限且自私的政治家还是愿意放弃实施她的意识形态，转而实施更温和的政策来维持政党的名誉。这样，就增加了这个政党在下一个阶段连任的机会，从而实施更符合她口味的政策。

15.2 多人合作博弈

13 和 14 章的分析已经证明当三个条件满足时，合作就可以继续维持：

> 第一，参与者对未来的收益足够重视；
> 第二，合作的频率足够高；
> 第三，有关过去行为的信息是公开的。

在这些条件下，每个参与者都不会选择违背合作准则，因为其他的参与者（一旦知晓这个背叛）将会在未来的合作中对背叛者采取有攻击力的报复。

在本节中，通过探索这些条件在一个大的群体且不那么友好的环境中是否有效来测试合作弹性。把群体定义为“大的”，并不是因为群体数量本身，尽管在我们的许多例子中，群体都有成千甚至上万的成员，而是因为策略的两个特征相互作用。第一个特征是相同的两个人再次博弈的机会很小。例如，eBay 网的买家和卖家在未来可能不会再合作。这种情况和两个在一起工作的同事或是姐妹之间平日为争洗澡间的战争有明显的不同。第二个特征是其他参与者对在一个参与者身上发生的一切一无所知。所以，集市上的单个买家和卖家知道他们是如何交易的，但集市上的其他人不可能知道他们之间的交易情况。

即使参与者的生命是有限的，他们也看重未来的收益，但当群体结构如刚才所述时，合作能不能稳固保持还有待分析。如果两个人相遇的机会很小，合作不可能维持：如果相互合作的时间还很遥远，那么在未来合作中报复的威胁就不具有威慑力。当然，惩罚不可能针对受害人：原则上，背叛者可能被任何其他参与者在未来合作中惩罚。然而，这个惩罚机制有两个挑战：

> 第一，发现背叛者的参与人一定要找到对自己最有利的惩罚方式，即使他对她并没有做错什么；
> 第二，她必须知道与她博弈的、有欺骗行为的背叛者过去的情况，从而有权采取惩罚措施。

但如果背叛者的信息不为群体中其他人所知的话，那么这个人很可能“逍遥法外”。

我们对这些问题的探讨是在市场这个大背景下。第一个例子是 eBay 网上的虚拟市场。在上百万的买家和卖家中，两个人再次合作的机会很渺茫。eBay 清楚地意识到可能存在欺骗行为（卖家兜售劣质产品，或者是买家不支付货款）。为了其他参与者的利益，eBay 设计了参与者可以就他的买卖经历留言这个系统。对 eBay 这个简单的系统来说，反馈机制保证了合作的持续性。当然，现实可能比我们的模型复杂一点，实际上，在 eBay 这个平台上，我们将讨论反馈的真实性。

最后一点凸显了如何激励一个人给其他参与者提供真实信息的重要性。为了分析这个问题，我们仍然以市场为背景，但是把时间往后推。这里，我们考虑分析十三世纪欧洲的实体市场。在那些市场，一个令人惊叹的被称为“商人法”（Law Merchant）的机制通过报告不

良交易行为，为其他商人提供信息。尽管报告的成本很高，而且受益的是其他商人。

尽管在过去的七个世纪中，市场的种种细节都改变了，但基本模式不变：当两个商人再也没有博弈的可能时，如何激励出他们的诚实行为？

◯案例研究

eBay

人们交易最流行的平台之一是eBay。任何时候都有上百万的卖家或者通过拍卖或者标上"立即购买"来甩卖货物。而市场的另一方，上百万的买家登录eBay网页，完成交易。这是一个滋生欺骗行为的场所。卖家可能吹嘘已经用过的产品是新的，或者买家在拍卖会上中标却不支付货款。因为买家和卖家在未来合作的机会很小，惩罚欺骗行为的可能性就很小。eBay意识到这个潜在的矛盾，构建了允许买家和卖家同全体参与者分享信息的反馈机制。这个机制使商人在未来的交易中避免了同有不良交易行为的商人合作，从而惩罚背叛者。

让我们从eBay的简单模式开始，分析之后讨论实际中存在的复杂情况。在每个阶段，买家和卖家随机买卖。可以想象一个买家在众多卖家中寻找卖他所需要产品的卖家。在他们的交易中，卖家选择商品的真实质量，它可以是优，一般，或者伪劣产品。买家不可能知道产品的真伪，除非在做出交易之后。如表15—3所示，卖家的成本随着质量增加，买家所得的价值和货物的好坏相关。一个优良产品，卖家的成本为13，对买家来说值30；一个一般的产品，卖家的成本为8，对买家来说值15；一个伪劣产品，卖家的成本只有2，但对买家来说没有任何价值。

表15—3 与产品质量有关的成本和价值

真实质量	实价的成本	对买家的价值
优	13	30
一般	8	15
伪劣产品	2	0

选择商品的真实质量后，卖家在eBay网上标价。为了简便，假设仅仅有三种价格：5，10和20。（我们也可以使卖家在网上对商品的质量进行报告。）报告可能是真的或者是假的，但是这也只会使分析变得复杂。在决定是否"立即购买"之前，买家会浏览商品价格和卖家的反馈评分。反馈评分是由曾经和卖家交易过的买家留下的正面或负面评价。

如果买家选择不买商品，那么她的收益为0。如果买了，那么她的收益为附加在商品上的真正价值（只有在买了之后才能知道），真正价值小于卖家的标价。假设卖家规定只有买了产品之后才知道产品的价值；那么如果这个产品没卖出去，他的收益就为0。表15—4说明了卖家和买家的收益取决于产品的价格和质量。在交易完成后，买家留下反馈信息。和在eBay网实际情况相反，我们不允许卖家留关于买家的信息。

表15—4 价格和质量对收益的影响

真正的质量	价格	卖家的成本	卖家的收益	买家获得的价值	买家的收益
优	20	13	7	30	10
优	10	13	−3	30	20
优	5	13	−8	30	25
一般	20	8	12	15	−5
一般	10	8	2	15	5
一般	5	8	−3	15	10
伪劣产品	20	2	18	0	−20
伪劣产品	10	2	8	0	−10
伪劣产品	5	2	3	0	−5

合作的阶段是无限的，但相同的买家和卖家却不一定再合作。卖家在每个阶段都期望遇见不同的买家。当一个买家浏览卖家的网页时，她所知道的卖家全部信息就是反馈评分。假设卖家的收益为他利益链上的现值，且每个买家只和一个卖家有一次交易。要记得，当她不买产品时，获得的收益为0，所以只有当她能得到正的收益时（这意味着产品的价值超过它的价格），她才会购买产品。

思考一下策略集合：

- 卖家的策略：

  如果卖家网页没有负面评价，选择以20的价格卖质量为优的产品。

  如果卖家网页有一条负面评价，选择以10的价格卖质量为一般的产品。

  如果卖家网页有两条或更多的负面评价，选

择以 5 的价格卖伪劣品。

- 买家的策略：

  如果卖家网页没有负面评价，就买。

  如果卖家网页有一条负面评价，如果价格为 5 或 10 就买，如果为 20 就不买。

  如果卖家网页有两条或更多的负面评价，就不买。

- 买家的反馈策略（假设她已买了产品）：

  如果（1）产品的质量为优，或（2）产品的质量一般，价格为 5 或 10，给予正面评价；

  如果（1）产品的质量一般，价格为 20，或（2）产品是伪劣产品，给予负面评价。

在卖家的策略下，买家估计到从没有负面评价的卖家买到的是质量为优的产品，从只有一条负面评价的卖家买到的是质量一般的产品，从有两条或更多负面评价的卖家买到的产品是伪劣产品。从这些信息中，让我们评估买家的最优策略。

如果卖家网页没有负面评价，那么按照买家的策略，她就应该买。即使价格为 20，她的期望收益仍然为 10，因为她估计得到的是质量为优的产品，价值为 30。很明显，如果价格越低，她的收益就越高。如果卖家有一条负面评价，那么只有产品价格为 5 或 10 时，买家应该买。因为价格如果为 20，她的收益就为−5，因为她从这样的卖家只可能买到质量一般而不是质量为优的产品，而质量一般的产品价值仅为 15。但如果它的价格为 5（或者 10），那么买家的收益为 10（或者 5），那么买家就倾向购买。最后，如果卖家有两条或更多负面评价，那么买家只可能得到无价值的伪劣产品，很明显，即使价格为 5，买家的最优策略也是不买。所以，买家的购买策略是最优的。

至于她的反馈策略也略占优势，因为我们假定留下反馈信息没有任何成本，买家的未来收益也和她自己留下的反馈信息没有关联。提供真实的反馈信息（这也是我们假设的）是最优策略。当然，留下假的反馈信息也是最优策略，比如，在网上留言说物流很慢而实际上商品按时到达了。然而，我们关注的焦点是找到一个均衡。在这个均衡中信息是可靠的。我们将来会回到这一点来讨论 eBay 网上交易的实际情况。

接下来，让我们预算卖家的最优策略。我们从卖家网页有两条或者两条以上的负面信息开始研究。那样的话，不管他怎么做都无关紧要，因为没有买家会买他的东西，他很可能会提供伪劣产品。不管他选的质量如何，他的收益还是为 0，在这种情况下，他提供的是伪劣产品，按照卖家策略的规定，这个决定和其他的决定的结果一样。

如果卖家网页有一条负面评价，那么买家估计他提供的是质量一般的产品。卖家的策略使他提供价格为 10 的质量一般的产品。根据买家策略，这样产生 2 个收益。因为买家拍单，而且留下正面反馈信息。所以，卖家仍然只有一条负面评价，这意味着他可以期望在未来从利益链上得到 2 个收益。卖家从价格为 10 且质量一般的产品中得到的收益为 $2/(1-\delta)$。而标价为 20 的话，他知道肯定没有人来买，因为买家不会为质量一般的产品付 20 个价格，所以卖家的收益为 0。因为未来的利益链仍为 2，标价为 20（没成交）的策略就劣于标价为 10（成交）。标价为 10 的策略也优于标价为 5 的策略，因为更低的标价意味着更低的收益。标价为 10 质量一般的替换策略就是提供价格为 10 的伪劣品，尽管他自己并不想提供价格为 5 或 20 的伪劣品。那么现在的收益就为 8，这比提供质量一般的产品获得的收益高。然而，一旦买家发现产品是伪劣的，她就会留下负面信息。因为卖家现在有两条负面信息，所有的买家在未来交易前都估计他提供的是伪劣产品，所以不会买他的产品了。他未来的利益链就为 0。总之，当：

$$\frac{2}{1-\delta} \geqslant 8 \Rightarrow \delta \geqslant \frac{3}{4} \tag{15.7}$$

时，提供伪劣品的策略就不如提供质量一般产品的策略。

最后，思考卖家网页没有任何负面评价的情况。他的策略使他提供价格为 20 的质量为优的产品。这就使得现在和未来的利益链为 7，获得的收益为 $7/(1-\delta)$。他也可以提供一般质量的产品，如果他以标价为 10 提供了一般产品，那么现在的收益就降低到 2，而未来的利益链不变（因为买家会给他正面反馈信息）。所以，这样做没有好处。或者他可以坚持为质量更低的产品标价 20。因为不管质量一般还是伪劣产品，他都会得到一条负面评价，这样他肯定提供伪劣产品，因为这使他现在的收益更高。以价格 20 卖伪劣品，他的现在收益为 18，而由于这一条负面评价未来的利益链只有 2。所以，当：

$$\frac{7}{1-\delta} \geqslant 18+\delta\left(\frac{2}{1-\delta}\right) \Rightarrow \delta \geqslant \frac{11}{16} \tag{15.8}$$

时，没有负面信息的卖家的最优策略就是以20的价格提供质量一般的产品。总之，如果表达式（15.7）和表达式（15.8）都成立的话，或者当$\delta \geqslant \frac{3}{4}$时，这个策略组合就是一个均衡策略。

信息反馈在eBay网的作用就是使买家对卖家已往的交易有一个了解。这就提供了一个激励卖家为买家提供高质量产品的机制，因为如果他不这么做，那么买家就会向全世界证明他的不良行为。后果就是未来的买家在未来认为这个卖家提供的是低质量的产品，卖家的获利交易就会减少（如果他还有顾客的话）。所以，未来的买家会惩罚卖家的欺骗行为，而正是惩罚机制阻止了卖家欺骗买家的行为。

在现实生活中，反馈机制运行良好，尽管也有一些弊病。我们的模型和现实的一个重要的区别就是在eBay网上允许卖家和买家都提供反馈信息。而且，评论不是同步进行的。例如，一个卖家在知道买家如何评论他之后才提供关于买家的信息，结果可能导致一个交易者即使被敲了竹杠还说另一个交易者的好话。原因如下：通过提供反馈信息，一个人为其他人提供有用的信息，他不是在帮自己。如果一个卖家欺骗了买家，这个买家知道了就再也不买这个卖家的产品。通过提供负面评价，他使其他没有意识到这个卖家不良行为的其他买家受益。对提供反馈信息的人来说，没有任何利益可图。在我们的模型中，提供信息也没有任何成本，所以最优策略就是提供可靠的信息。但现在假设卖家会通过给买家负面评价报复买家提供负面反馈的行为。这就降低了买家的反馈评分，从而影响她未来的交易。所以，提供虚假信息对提供者没有好处，但还有潜在的成本。这种不对称性使得那些不管产品质量如何都提供正面信息的人产生偏见。

如表15—5所示，一个近来的研究发现，留下的评论中，99%是正面信息，[①]而大部分对交易满意的商人可能只是简单地描述交易的情况。再仔细检查，我们就会发现比报复性评论更邪恶的例子。当买家留下正面信息，卖家有99.8%的可能会回馈正面信息。然而，当买家留下非正面信息，买家只有60.7%的可能回馈正面信息。如果仔细研究留下的实际信息（不仅仅是关注记录的信息是正面的、中立的还是负面的），报复的丑陋痕迹就展露无疑。例如，思考关于购买iPod的交易：

买家：产品的操作性不好。得不到卖家的回复。

卖家：不是诚实的eBay会员。非常差的交流沟通能力。

表15—5　eBay网上的反馈信息（%）

评论类型	买家对卖家的评论	卖家对买家的评论
正面	51.2	59.5
中立	0.2	0.2
负面	0.3	1.0
不做评论	48.3	39.4

正如这个调查所显示的，一些重要的东西在我们简单的eBay模型中被掩盖了。在下个例子中，我们将首先解决这个问题，尽管不是在网络上，而是在13世纪的欧洲市场。不再假设提供信息是无成本的，我们假设对一个商人来说是有成本的，而且，我们也不指望他的信息是真实的。在中世纪的环境中，市场机制使得有些商人被诱惑提供虚假交易信息，而他们的邪恶行为是由不公正的判决导致的。



点拨

即使两个人在未来并没有合作的可能，如果他们的交易信息暴露给群体其他参与者，那么合作还可以维持。

15.2 理解测验

现在假设卖家只要得到一条负面信息就提供伪劣产品（价格为5）。像以前一样，当他没有任何负面信息时他选择提供价格为20的优质产品。假设买家使用和以前相同的反馈策略。找到买家的最优购买策略，推导保证卖家策略最优的条件式。

中世纪的商人法

在欧洲十三世纪的市场上，一个商人到一个村庄卖

① 从表15—5可知，交易后有51.2%给出正面评价，51.7%给出评价（或者是正面，或者是负面，或者评价不好不坏）。0.99=0.512/ 0.517，意思是说，当有第一个评价时，这条评价99%是正面评价。——作者注

东西。一个买家看了产品试图估计它的质量，而卖家打量买家看他是否能及时付款。如果双方都对彼此信任，相信他们的讨价还价可以持续到最后，那么交易就发生了，否则，交易就不可能发生。

如果这对买家和卖家期望未来可能的合作，双方的聪明做法就是做到公平来维持名誉。但是假设这种未来相遇的机会很渺茫，而且买家和卖家都不指望在未来的任何时间进行另一笔交易。这样他们还有相信彼此的可能吗？未来的合作还会发生吗？

为了解决这些问题，一个名为“商人法”的机构在欧洲兴起，“商人法”管辖欧洲许多商业交易，就像一个法官，裁决遭受损失的一方可以采取报复措施。



中世纪的市场

尽管商人法可以进行裁决，但它没有法律效力来保证损失得到赔偿。既然没有法律效力来保证裁决得到实施，那么商人法是如何在提高诚信贸易中起作用的呢？

商人 2

商人 1		诚实	不诚实
	诚实	1, 1	-1, 2
	不诚实	2,-1	0, 0

图 15—3　交易中的博弈

为了解决这个问题，思考下面的博弈：每个商人的生命都是有限的，在每个阶段他都有机会和其他任何人做交易。如果两个商人中的任何一个决定不参与交易，那么每个人获得的收益都为 0。如果他们决定交易，那么这场交易就和“囚徒困境”的情况相似。如图 15—3 所示，商人同时决定是采取诚实还是不诚实策略。

在他们做出决定之后，每个人得到他们的收益然后两个人分道扬镳。尽管每个商人都希望有源源不断的交易机会，他和同一个商人再次交易的机会还是很渺茫。所以，尽管商人一直在交易中，但每次的交易都是一个独特的例子。假设一个商人的收益为他货币流的现值，这里他的折现系数为 δ。

当两个商人相遇，每个商人都知道自己过去的交易经历，但不知道另一个商人过去的交易行为。所以，一个欺骗者不可能在未来被惩罚，因为未来和他交易的搭档对他过去的行为一无所知。那么在这场博弈中，只存在一个子博弈精炼纳什均衡，它使两个商人要么不交易，要么在交易中采取不诚实的策略。这两种情况都导致收益为 0，而如果他们都在交易中表现诚实，那么每个商人的收益为 1。

现在我们引入商人法。在每个阶段，每次商人交易有六个步骤：

- 步骤 1（咨询）：每个商人决定是否去向商人法咨询了解另一个商人是否存在欺骗行为。而这样做的话，商人需要付款，价格为 p。（商人法是盈利机构，而不是造福社会的机构。）
- 步骤 2（交易）：在了解了另一个商人的信息之后，两个商人决定是否达成交易。如果他们决定达成交易，那么他们按照表 15—3 的规则交易。
- 步骤 3（申诉）：交易结束之后，如果觉得自己被骗，任何一个商人都可以去商人法那里请求裁决。原告（申诉的人）向商人法寻求裁决的价格为 q。
- 步骤 4（裁决）：如果商人申诉，商人法就应当做出正确裁决。如果一个商人诚实而另一个商人不诚实，那么商人法就裁决第一个商人给第二个赔付 d。任何其他结果，不得索赔损失。
- 步骤 5（付款）：如果其中的一个商人被商人法告知赔偿损失，他决定是否做出赔偿。
- 步骤 6（记录）：如果裁决赔偿损失，而有罪的一方不予赔偿，那么这个案子就被商人

法记录在案，成为商人法永久记录的一部分。（如在步骤 1 中描述的，这条记录可以供任何人以价格 p 查询）

以下策略描述了所有商人在个人交易历史的基础上，在每个步骤中的做法。

- 步骤 1（咨询）：如果一个商人没有不良记录，那么他会向商人法咨询另一个人的记录。如果一个商人自己有不良记录，那么他不会向商人法咨询另一个人的记录。
- 步骤 2（交易）：如果两个商人都不去咨询或者咨询表明至少有一个商人有不良记录，那么他就不会做出交易行为。如果两个商人都去咨询且两者都没有不良记录，他就做出诚实交易行为。
- 步骤 3（申诉）：如果交易达成，一方诚实，另一方不诚实，那么受害者把案件上诉到商人法，而欺骗者不会上诉。
- 步骤 5（付款）：如果案件被上诉到商人法且商人法发现被告有罪，责令其赔偿 d，如果被告以前没有不良记录他就会赔付受害者的损失，如果有，他就不予赔付。

在决定什么条件下策略是对称子博弈精炼纳什均衡之前，让我们首先了解一下如果所有的商人采取上述策略的行为意味着什么。假如所有的商人一开始都没有未付赔偿的记录，也就是商人法那里没有不良记录登记。注意，这并不是说一个商人会永远诚实做交易，但如果他在交易中欺骗了别人，那么商人法就会裁决他赔付损失。按照自己的策略，一个没有不良记录的商人会去商人法那里了解其他商人的情况。每个商人会了解另一个商人有无未付赔偿记录，这意味着他们会本着诚实的原则做买卖。如果是这样的话，两个商人都不会去商人法那里申诉。所以，如果商人交易前没有未付赔偿记录，那么交易后也不会有未付赔偿记录。他们在这个阶段中的收益为 $1-p$，因为从诚实交易中得到 1，而付给商人法 p。因为每个商人都期望每个阶段都是这样的结果（尽管每次是和不同的商人做交易），他的货币流的现值就为 $(1-p)/(1-\delta)$。如果其他商人都采取同样的策略，这就是每个商人从这个策略中得到的收益。

为了证明这个策略是最优的，我们需要证明对有任何交易历史的商人来说，没有其他的策略产生更高的收益。证明这一点需要考虑每一个步骤规定的行为，在每个步骤中商人必须做出某种决定。

什么情况下咨询商人法是最优策略？首先注意一个商人的未来收益与他是否去咨询商人法无关。如果他没有未付赔偿记录，那么他的策略就使他不去欺骗其他人，不论他去没去咨询商人法；所以，在这个阶段后，他没有不良记录。如果他有未付赔偿记录，那么这个记录就是一个永远的伤害，这种情况下，未来他很可能没有任何交易的机会，不管他是否去咨询商人法。

于是，咨询决定的最优策略取决于它如何影响一个商人当前的收益。假如他没有未付赔偿记录。如果他咨询了商人法，那么他期望了解到的另一个商人也没有不良记录，这种情况下，他们会诚实地达成交易得到当前收益 $1-p$，这里我们已扣除咨询的成本。如果他不咨询商人法，那么根据策略组合，不会有交易达成，所以他当前收益为 0。这样，当 $1-p \geqslant 0$ 或 $p \leqslant 1$ 时，最优策略就是去商人法那里询问。换句话说，当咨询费用不是很高的时候，最优策略是去商人法那里咨询。

现在假设他有一条未付赔偿记录。那么他从咨询中获得的当前收益为 $-p$，因为他必须付咨询费，而且其他商人在商人法处获得他有不良信息记录，就不会和他达成交易。如果不去咨询，获得的当前收益为 0。所以，如果一个商人有一条未付赔偿记录，他的最优策略就是不去商人法处咨询。

什么样的条件下诚实交易是最优策略？如果一个或者两个商人都没有去咨询商人法，另一种情况是至少有一个商人有不良记录，那么交易不会发生。所以，假设两者都去咨询商人法并且两者都没有未付不良记录。这样，从诚实交易中得到的收益为：

$$1+\frac{\delta(1-p)}{(1-\delta)}$$

一个商人得到当前收益 1（咨询费过去已经付了），凭着良好的记录，获得未来收益流 $1-p$。但如若不诚实交易（以后都这样做），那么这个商人的收益为：

$$2-d+\frac{\delta(1-p)}{(1-\delta)}$$

他从欺骗另一个商人那里得到的收益为 2，而另一个被骗的商人会去商人法那里申诉，得到赔偿 d。不诚实商人的策略使他赔偿那些损失，所以他从不诚实交易中获

得的当前收益为 $2-d$。因为他付了赔偿损失，所以他的未来收益流是 $1-p$。那么当：

$$1+\frac{\delta(1-p)}{(1-\delta)}\geqslant 2-d+\frac{\delta(1-p)}{(1-\delta)}\Rightarrow d\geqslant 1$$

时，诚实交易为最优策略。所以，赔偿损失必须足够多使得商人必须诚实交易。

什么样的情况下申诉是最优策略呢？假设一个商人欺骗了另一个商人，后者的策略使他去商人法那里申诉。这样做他花费 q，但他可以从裁决中获得另一个商人（只要他没有未付赔偿不良记录）的赔偿损失。当 $d-q\geqslant 0$ 时，就是说，当得到的赔偿至少和他付给商人法的费用一样多时，申诉是最优策略。

什么样的情况下支付赔偿损失是最优策略呢？这是最耐人寻味的例子，因为他涉及商人法是否可以不借助法律效力来引导商人赔偿损失。如果对一个商人（当前无任何未付赔偿损失不良记录）做出赔付裁决，那么赔偿后他的收益为：

$$-d+\frac{\delta(1-p)}{(1-\delta)}$$

他必须付赔偿损失，因为他没有任何不良记录，这样可以获得未来收益 $1-p$。相反，如果他不支付赔偿损失，那么他当前的收益为 0，因为有了一条不良记录，未来可能不会有任何交易机会。所以，当：

$$-d+\frac{\delta(1-p)}{(1-\delta)}\geqslant 0\Rightarrow\frac{\delta(1-p)}{(1-\delta)}\geqslant d$$

时，商人赔偿损失是最优策略。所以，赔偿额不能太大。

总结上述所有的分析结果，我们得出必须满足下列四个条件才能成为均衡策略：

- 当 $p\leqslant 1$ 时，咨询商人法是最优策略。
- 当 $d\geqslant 1$ 时，诚实交易是最优策略。
- 当 $d\geqslant q$ 时，向商人法申诉是最优策略。
- 当 $d\leqslant\frac{\delta(1-p)}{(1-\delta)}$ 时，赔偿损失是最优策略。

第一个条件表示咨询费不超过诚实交易获得的值。第二个条件表示赔偿损失足够多从而阻止商人有欺骗行为的发生。第三个条件表示当一个商人被骗时，申诉商人法的费用不要高的使商人觉得不值得去申诉。最后一个条件表示为了在商人法那里维持良好的信誉，赔偿损失足够低使得商人愿意支付赔偿。那么，损失赔偿额不能太高也不能太低。

尽管商人法没有法律效力，但作为信息的公开机构它在合作交易的促成中起着很重要的作用。商人们发现利用商人法的信息是最优策略（即使它不是免费的），而且这样做可以引导诚实交易。如果诚实交易并没有发生，商人法为了避免市场上其他人受骗，会将不良信息公之于众。

本章总结与练习题

本章总结

尽管在许多策略组合（如，行动事件，买卖车辆，法庭审理和战场）中的主导力量是斗争，但现实生活中还是存在很多合作的例子：人们互相帮助从而也帮助了自己。解释人们为什么合作并不需要假设人们是否出于关心：自私的人如果知道不合作就会损失未来合作的可能（不管是暂时的还是永久的），他们就会与别人合作。只要未来的惩罚足以对现在产生很大的负面影响，合作就引导出人们好的行为。

在这章里，我们已经证明这个机制不仅存在于期望未来有无数合作机会的两个人中，也存在于希望合作于不久后终止的多人合作中。在**世代交替**的模型中，每个人只在特定的阶段存在。因为她的生命即将终结，所以未来的利益对她的诱惑越来越小。在她生命的最后阶段，没有未来的利益也就没有任何办法引导她积极合作。然而，对群体的某些成员来说，合作是最优策略。思考某些人在群体中存在的倒数第二个阶段，如果这个成员采取背叛的策略，周围的其他人就会在未来的博弈中惩罚他。如果这个成员表现良好，周围

的其他人就积极合作在未来给他补偿。引导他积极合作的因素是群体的新补充的成员对他未来的惩罚。每一代人都被引导积极合作使得下一代人也与他们一样采取积极合作的行为，如此循环下去。维持合作的关键因素是群体的无限存在性。随着补充人员的加入，总有人来保证社会中的积极行为得到奖励，背叛行为得到惩罚。

第二个经典的例子是当群体规模很大时，两人再次遇见的机会就很渺茫。在这种情况下，尽管个人的生命足够长，两人再次合作的可能还是微乎其微。所以，这种合作是一次性的，每个人都面临着一次又一次这样的暂时性合作。为了使合作良好地进行，对不良行为的惩罚必然来自于与他未来博弈的人，因为他并不一定再次和他欺骗过的参与者博弈。关键是社会了解到这一情况，这就要求人们交换信息。所以，整个社会强制良好行为发生：如果一个人选择背叛，他就可能被未来的博弈者惩罚。我们知道 13 世纪欧洲的商人法和 21 世纪 eBay 网的反馈信息机制都在信息共享和提高合作中起了很重要的作用。

练习题

1. 在每个阶段，101 和 102 的房客博弈的情况如图 15—4 所示。房子无限存在，房客来了又走，每个房客只停留两个阶段。并且，当 101 的房客停留第一个阶段，102 的房客停留在第二个也是他的最后一个阶段；同样，当 102 的房客停留的是第一个阶段，而 101 的房客停留的是第二个也是她的最后一个阶段。每个人的收益都是在这两个阶段的单个阶段的收益之和。

		102 房客		
		x	*y*	*z*
101 房客	*a*	8,8	0,4	5,10
	b	4,0	3,3	8,1
	c	10,5	1,8	0,0

图 15—4　房间博弈

a. 求一个子博弈精炼纳什均衡使得一系列的博弈在 (c,x) 和 (a,z) 之间替换。

b. 求一个子博弈精炼纳什均衡使得每个阶段在 (a,x) 中博弈。

2. 有 20 个成员的俱乐部中的每个成员都在这个俱乐部停留 5 个阶段。在任何时候，有 4 个刚进入组织的，4 个刚进入第一个阶段的，4 个刚进入第二个阶段的，4 个刚进入第三个阶段的和 4 个刚进入第四个阶段的。组织有一个领导者，她的岗位是没人觊觎的、吃力不讨好的职位。个人成为领导者的成本为 x，而成为俱乐部成员得到的值为 v。假设 $x > v > 0$。不是领导者的成员得到的收益为 v，而领导者得到的收益为 $v-x$。每个成员的收益为在各个阶段收益的总和。（就是说，折现系数等于 1）。因为没有人想成为领导者，俱乐部的程序就为：随机选择某个人担任这个职务。一个领导者在职的时间只有一个阶段，所以随机选择应该在每个阶段开始时进行。在这场博弈中，一个人做的唯一决定就是这个人在被选中之后决定是否要担任领导者这个职位。如果他或她在被选中之后拒绝成为领导者，这个俱乐部就解散，这样的话，每个人得到的收益为 0。

a. 假设随机选择发生在那些已经在组织中停留了两个阶段的成员中，那么这个被选中的成员就在第三个阶段成为领导者。推导一个人被选中且承担这项工作的博弈成为子博弈精炼纳什均衡的条件。

b. 假设加入组织的会员费为 p。思考一种策略组合，在这种策略组合下，一个人加入组织并在第三阶段被选为领导者且承担职务。推导这个策略组合成为子博弈精炼纳什均衡的条件。

3. 返回到 15.1 小节中的部落防御博弈。思考使用同样的策略组合，但现在假设惩罚持续两个阶段，而非一个阶段。推导这个策略组合成为子博弈精炼纳什均衡的条件。

4. 思考在世代交替群体中，在任何时候，都有三代人，每个人生活三个阶段。在每个阶段，他们同时在三

种行动中做决定：低，中和高。这场博弈是个对称博弈，每个人的收益如表 15—6 所示。比如，如果一个参与者选择高，一个选择中，一个选择低，那么他们的收益分别是 0，2 和 5。如果没有人选高，两个参与者选择中，那么选择中的每个参与者得到收益为 1，选择低的参与者得到收益为 4。

假设合作博弈中，一个参与者在他生命中的第一个阶段选择高，在第二个阶段选择中，在最后一个阶段选择低。思考以下策略组合：如果这场博弈中合作一直持续，那么一个参与者在他的第一个阶段选择高，第二个阶段选择中，在人生的最后一个阶段选择低。另一种情况是不管他处于哪个阶段，都选择低。推导这个策略组合成为子博弈精炼纳什均衡的条件表达式。

表 15—6

选择高的人数	选择中的人数	选择…参与者的支付收益		
		…高	…中	…低
0	0	x	x	0
0	1	x	−1	2
0	2	x	1	4
0	3	x	3	x
1	0	−2	x	3
1	1	0	2	5
1	2	2	4	x
2	0	1	x	6
2	1	3	5	x
3	0	4	x	x

5. 在电影《动物屋》(*Animal House*) 中，奇普·迪勒 (Chip Diller) 正宣誓进入 Omega 社团。依进入社团的仪式，奇普脱下内裤，弯下腰，扣住他的脚踝，而此时高年级兄弟会主席用木棍打奇普的后臀，奇普回应到“谢谢你，先生。可以再来一遍吗？”为了便于分析，我们假设学生在大二学期开始宣誓加入兄弟会，如果他们要加入 Omega 社团，必须要被打屁股且得经得住打。假设宣誓时被打屁股的值为 m，这里，对大多数人说，$m < 0$（就是说，宣誓加入社团是要付出代价的）。成为 Omega 社团成员每年得到的值为 v，且 $v > 0$。高年级的成员应该参加这种欺辱别人的行动，值为 s。s 可以是正值，如果你从中得到乐趣；也可以是负值，如果你讨厌做这个的话。每年的收益折现率为 δ。比如，如果奇普加入 Omega 社团，作为大二的学生，他参加这种欺辱行动（作为受痛者），作为大四的学生，也参加这种欺辱行动（作为施痛者），他得到的收益就为 $(v+m)+\delta\times v+\delta^2\times(v+s)$。假设没有加入社团的同学得到的收益为 0。这场博弈是个世代交替的博弈。在任何时候，都有大二学生，大三学生和大四学生；然而，只有大二和大四学生做决定。大二学生决定是否要被欺辱获得成为社团成员的资格，不经历这场侮辱就不能加入社团。大四学生必须决定是否要参加欺辱别人的行动，如果不参加就被取消成员资格。思考一个学生在大二和大四都参加这场欺辱行动的对称策略组合。

a. 推导这个对称策略组合成为子博弈精炼纳什均衡的条件式。

b. 假设 $v=0$ 且 $m < 0$。证明只有那些施虐狂才会加入社团。

c. 假设 $v=0$ 且 $\delta=0$。证明只有那些施虐或受虐狂才会加入社团。

6. 有一个大的群体，在每个阶段所有参与者随机搭配进行图 15—5 的对称性博弈。没有两个参与者有第二次博弈的机会。每个参与者的收益为单个阶段的收益之和，这里折现系数为 δ。

假设每个参与者过去的行为为全体成员所知。求在每个阶段中 (c,c) 为最终结果的子博弈精炼纳什均衡的条件表达式。

		参与者 2		
		a	*b*	*c*
参与者 1	*a*	1,1	0,2	7,0
	b	2,0	3,3	4,1
	c	0,7	1,4	5,5

图 15—5

7. 让我们思考 15.2 小节中 eBay 网的另一种游戏规则。假设 eBay 网上买家为卖家评论有时间限制：在一个阶段（一次交易）的末尾和另一个阶段（下一次交易）的开始之间。并且，一个买家由于技术方面的原因可能不能在那个时间段对卖家评论。所以，所有的买家和卖家都知道有一个概率 r（$0 < r < 1$）买家不能对卖家评论。思考 15.2 小节中的策略组合。推导这个策略组合成为子博弈精炼纳什均衡的条件表达式。

8. 思考群体中有装修公司和房主。在每个阶段，一个房主和一个装修公司博弈。房主决定是否雇用这个公

司来改造房屋。如果她选择这个公司，那么由公司决定是提供高质量的工作还是低质量的工作。相关收益在表15—7中给出。一个房主如果不对房屋进行改造，得到的收益为10，而一个装修公司如果当前没被雇用得到的收益为0。

表15—7 装修博弈的收益

工作质量	装修公司	房主
低	30	5
高	20	25

装修公司永久存在，他们的收益为各个阶段的现值总和，这里折现系数为δ。房主经常雇用房屋改造公司。

a. 假设房主了解装修公司过去所有的工作情况。求一个子博弈精炼纳什均衡，在这个均衡中，结果是房主雇用装修公司，且使他提供高质量的工作。

b. 假设一个房主只了解了一个装修公司过去有多少顾客，且如果他有一些顾客，他就能了解到由这个公司给房主提供的顾客的改造情况。证明子博弈精炼纳什均衡必须导致所有的房主不雇用任何装修公司。

第六部分 演化博弈

第 16 章

演化博弈与生物学：演化稳定策略

学习目标

16.1 鹰鸽博弈
16.2 演化稳定策略
16.3 演化稳定策略的特征
16.4 多总体模型
16.5 敌意的演化

本章导读

我的进化理论是以达尔文的演化理论为基础的。

——史蒂芬·赖特（Steven Wright）

在这本书中，我们已经用博弈理论来阐述人们在采取策略时的行为。这类研究已被用于确认人们使用的策略集，这些策略集具有这些特征，即人们根据他人的行为如何来提高自身的生存状态。在这章和接下来的章节中，我们试图用这些方法来理解其他动物的行为——不仅是其他的灵长类动物如大猩猩，还包括鸟类、爬行动物、鱼类甚至昆虫类。实际上，我们试图找出方法可以使我们理解任何星球上的任何生物（让我们大胆点吧）。这些方法可以解释动物交配仪式，对资源的竞争、合作、交际和下代生物的性别比例。这项计划用来探索我们前几章提到的无法进行有意识的思考的低智力的生命形态。当然，对蚂蚁而言理性不是共同知识！

如果用博弈论探索整个动物王国，我们依赖的这些机制选择的策略就不能有认知意义。正如达尔文借用社会科学家托马斯·马尔萨斯（Thomas Malthus）的观点来扩展《物种起源》一样，我们现在也需要借用达尔文的理论。选择策略组合的机制是自然选择的结果，在持续、随意的进化过程中，适者生存，不适者被淘汰。自然选择会导致一个稳定的策略选择机制。人类的智慧被“自然”选择的“智慧”所替代。

我们研究人类行为的方式和研究其他动物的行为的方式有极强的相似性。为了理清这一点，我们回到前面章节提到的方法，现在我称之为**完全信息博弈**。借助这个策略，我们以几个参与者为研究对象，每个都对博弈结果的收益函数有一定的了解。每个参与者可以使用一些策略，而在这些策略中，有一个成员采取的策略都对结果有决定性作用。一种策略组合（即

每个参与者选择的策略方式）决定了博弈的结果。每个参与者在选择策略时，应根据其他参与者的信息依靠理性来选择可以使自己收益达到最大的策略。纳什均衡是稳定策略组合的一个特点，纳什均衡使每个参与者根据其他参与者的信息最大化自己的收益。这个特点假定知道其他参与者选择的策略。

研究动物王国所用的方法称为**演化博弈理论**。研究切入点和策略式博弈不一样，演化博弈理论是设定在一个更大的情境中：群体参与者，每个参与者都是经过选择的。战略型博弈可能是双方的博弈（如两雄性为了繁衍后代争夺一雌性），随机的两个参与者完成两人博弈。然而，他们并不能选择策略，而策略是理性博弈理论的一个重要的出发点。相反，每个成员被赋予一种策略。所以，每个参与者被设定了参与的规则，不像在理性博弈理论中有自由意志的发挥。选择策略的过程在群体层面发生，而非个人层面。执行更好的策略（这里称为适应值而非收益）依据自然选择机制，将替代执行较差的策略。尽管每个参与者不能变更自己的行为，参与者采取策略的比例可以演化：更适合环境的参与者有更多的后代，继承了这些策略的后代更好地适应了环境。正如理性博弈一样，我们也要寻找一种稳定的结果，更具体地说，一系列的策略使群体随着时间一直生存繁衍下去。在演化过程中，群体是稳定的，对突变有很强的适应力。表 16—1 总结了理性博弈理论和演化博弈理论。

表 16—1　理性博弈理论和演化博弈理论

理性博弈理论	演化博弈理论
参与者集合	从参与者集合中选出一个群体
收益：生存状态值的量化	适应值：繁衍力的量化
参与者选择策略	从一个参与者那继承策略且由自然“选择”
均衡：每个参与者的收益达到最大值	均衡：少量的突变体不能在群体中存活

尽管博弈理论的这两个分支的方法有一些共性，但强调它们之间视角的不同却是至关重要的。理性博弈的分析对象是参与者，目的是描述每个参与者做了什么。在演化博弈理论中，分析单元是策略（特性或者基因）和我们试图描述的群体采取的混合策略。来来去去的群体参与者是谁无关紧要，重要的是群体所采取的混合策略。

这一章和下一章的主要目标是理解在动物王国中给定的战略环境下如何寻求稳定的混合策略。我们需要定义什么是“稳定”，然后学习如何寻求这样一种状态。在这章，我们的关注点主要是同类生物，所以所有的参与者采取同样的策略。大体上来说，一个群体的稳定策略，也就是**演化稳定策略（ESS）**，必须满足这个条件，即它必须能抵御一小部分的突变者群体，这类群体采取不同的策略。从自然选择的角度来说，演化稳定策略比突变策略更具竞争力。在 16.1 小节和 16.2 小节中，我们使用动物王国里的普通争斗模型——鹰鸽博弈来做研究。在给鹰鸽博弈和蜣螂（别怀疑你的眼睛，就是粪上活动的生物，你会喜欢这些东西的）的有趣争斗寻求一种演化稳定策略后，16.3 小节我们讨论演化稳定策略的特征，包括它和纳什均衡的关系。这章以思考基本结构的两种变形结束。在 16.4 小节，我们将会介绍多元群体可能涉及雌雄的战略环境。然后，在 16.5 小节，介绍演化稳定策略如何适应群体规模小的情况以及证明进化过程中敌对是如何产生的。为了更好地理解这章所讲的方法，你需要温习第 7 章提到的混合策略。

使群体更适合生存的策略淘汰那些不适合的策略很好地解释了演化论研究方法。在定义演化稳定策略的同时，我们也试图寻找确定从动态中演化出的群体结构，这些动态过程没有明确地被模型化。在演化稳定策略中，可替换的方法就是把这些动态过程模型化，我们在 17 章就是这样做的。使用的模型被称为**复制者动态**（replicator dynamic），从稳定群体方面来说，复制者动态产生的结果与演化稳定策略有很强的关联性。

16.1 鹰鸽博弈

攻击性是人类内在的、独立的、本能的特性，它包括对文化的强烈抵触。

——西格蒙德·弗洛伊德

在发情期，赤鹿常常处于争斗中，这使得它们变得丑陋。如果它们碰见资源（比如雌鹿或者一片领地），它们就会通过一项仪式来决定谁拥有资源。最经典的仪式就是咆哮。因为身体强壮者咆哮时才有能量，胜负可以在这个阶段揭晓。如果不能，它们可能通过锁住鹿角

或者推搡彼此进行身体搏斗。这场争斗可能导致有的鹿受伤，但也能决定谁是胜者。

生物学家把这种模型称为鹰鸽博弈。在鹰鸽博弈中，有两种可能的策略：鹰策略和鸽策略（哦，真是个惊喜！）。如果一只鹿采取鹰策略，它就主动出击获取资源。相反，一只采取鸽策略的鹿就会做出争斗的姿态（通过咆哮），但当其他动物也来争斗，它就会退缩。从适应值来说，假设资源值为 V。所以，如果一只采取鹰策略的鹿和一只采取鸽策略的鹿争斗，采取鹰策略的获得资源，因为当其他的鹿通过锁住鹿角的方式加入争斗时，采取鸽策略的鹿就会退缩放弃。采取鹰策略的鹿就获得适应值 V，而采取鸽策略的鹿获得适应值 0。如果两只鹿都采取鸽策略，它们会长时间处于做出争斗状态直到其中一只放弃争斗，灰溜溜走开为止。如果我们假定每一方都有 50% 的概率坚持的时间比另一方长，当两方争斗时，期望的适应值就是$\frac{V}{2}$。最后，如果两只采取鹰策略的鹿决斗，争斗持续直到其中一只受伤。失败者的适应值就是 $-C$（这里的 C 是受伤的成本），而争斗中的胜利者就获得了资源，得到适应值为 V。当两只都采取鹰策略的鹿一起博弈，它们获胜的概率一样大，那么它们的适应值就是（$V-C$）/2。

如图 16—1 所示，这些适应值给我们解释了战略型博弈。假定 $C>V>0$，那么受伤成本的适应值就大于得到领土利益的适应值，这种结果当然是这样的情况，即负伤的程度足以威胁生命。

赤鹿 2

		鹰	鸽
赤鹿 1	鹰	$\frac{V-C}{2}$，$\frac{V-C}{2}$	V，0
	鸽	0，V	$\frac{V}{2}$，$\frac{V}{2}$

图 16—1　鹰鸽博弈

在分析自然界的这种情况之前，我们先描述鹰鸽博弈在哪些地方和人类相似。假如不期而遇的两个人在价值为 V 的资产上讨价还价，一个人可以采取大胆的策略（鹰策略），在这种情况下，即使成本很大他也愿意去争取资产，他也可以采取小心谨慎的态度（鸽策略），这样他就可以避免高昂的谈判成本，即使这样做意味着他会失去很多。如果两个人都小心谨慎，他们能保证利益均分，每个人得到$\frac{V}{2}$。如果两个人都很大胆，他们会冒着失去这单资产的风险争斗。假定每个人的期望收益值少于$\frac{V}{2}$，比如（$V-C$）/2。如果一个人大胆，另一个人小心谨慎，那么后者放弃，因为他知道采取大胆策略的谈判家会不惜一切代价得到这笔资产。所以大胆的人就得到适应值 V，而小心谨慎的人就得到适应值 0。

现在我们返回动物王国和那些紧紧锁住鹿角的谈判对手。我们就以所有参与者都是被赋予鹰策略的群体做研究。群体中的争斗参与者随机匹配，因为双方都采取鹰策略，每个获得的适应值就是（$V-C$）/2。现在，假定群体中有一小部分突变体被赋予鸽策略，概率为 ε，所以有 $1-\varepsilon$ 的比例为采取鹰策略的参与者，采取鸽策略的比例为 ε，这里 ε 为正数，但很小。于是采取鹰策略得到的期望适应值就是：

$$(1-\varepsilon)\times\left(\frac{V-C}{2}\right)+\varepsilon\times V \qquad (16.1)$$

在概率 $1-\varepsilon$ 下，采取鹰策略的参与者和另一只采取鹰策略的参与者博弈，得到的适应值为（$V-C$）/2，而在概率 ε 下，它与另一只采取鸽策略的参与者博弈，得到适应值 V。那些突变体，即被赋予鸽策略的参与者得到的期望适应值就是：

$$(1-\varepsilon)\times 0+\varepsilon\times\frac{V}{2} \qquad (16.2)$$

这些突变体鸽遇见采取鹰策略的参与者概率为 $1-\varepsilon$，获得的适应值为 0，而遇见另一个采取鸽策略的参与者的概率为 ε，得到的适应值为 $V/2$。

对于只有采取鹰策略成员的比例稳定的群体来说，采取鹰策略的参与者一定会比采取鸽策略的突变体得到的适应值高。如果不是这样的话，突变体就不会被淘汰出局。采取鹰策略成员的比例就得满足这个条件表达式（16.1）>（16.2）：

$$(1-\varepsilon)\times\left(\frac{V-C}{2}\right)+\varepsilon\times V>\varepsilon\times\frac{V}{2}$$

解这个表达式我们得到关于 ε 的不等式：

$$\varepsilon>\frac{C-V}{C} \qquad (16.3)$$

既然（$C-V$）/C>0，当突变比例 ε 小于（$C-V$）/C 时，不等式（16.3）就不成立。

这个分析告诉我们的是当鸽突变体的比例足够小，它就可以成功侵入采取鹰策略的群体。当一个群体中几乎全是采取鹰策略的成员的时候，采取鹰策略的参与者遇见的采取鹰策略参与者的概率就大，这样他们博弈完

成后，每个参与者就有 50% 的可能负伤，50% 的可能获得对资源的控制权而又不受丁点伤。这样的结果并没有很大的吸引力，因为资源值 V 比负伤成本 C 小，鹰付出的代价就太大。现在，因为突变为鸽的比例很小，采取突变策略的参与者遇见采取鹰策略的参与者的概率就很大，在这种情况下，采取鸽策略的参与者就退而不战了。尽管它没有得到资源的控制权，它也没有任何损失。当采取突变策略的参与者遇见采取鹰策略的参与者时，突变体得到的适应值为 0，这大于采取鹰策略的参与者与采取鹰策略的参与者博弈得到的适应值 $(V-C)/2$，因为 $(V-C)/2<0$。现在，采取鹰策略的参与者也有遇见采取突变策略参与者的可能性，这样采取鹰策略的参与者得到的适应值就高，为 V；当采取突变策略的参与者遇见突变体时，得到的适应值就低，为$\frac{V}{2}$，但这种情况很少，因为鸽的比例很小（记住，鸽的突变比例被假定很小）。所以，决定采取策略适应值的是当它遇见随处可见的鹰时如何执行策略，在这种情况下，采取突变策略就优于采取鹰策略。

那么，只包含被赋予鹰策略的参与者的群体就不稳定了，因为它可以被小部分的突变体侵入。那么，只含有被赋予鸽策略的群体情况如何呢？它对突变适应吗？思考一下群体中突变为鹰的比例为 ε，那么采取鸽策略得到的期望适应值就是：

$$(1-\varepsilon)\times\left(\frac{V}{2}\right)+\varepsilon\times 0 \qquad (16.4)$$

因为它遇见另一只采取鸽策略的概率为 $1-\varepsilon$，在这种情况下，它获得的适应值为$\frac{V}{2}$，遇见突变体的概率为 ε，这样它获得的适应值为 0。突变体采取鹰策略的期望适应值为：

$$(1-\varepsilon)\times V+\varepsilon\times\left(\frac{V-C}{2}\right) \qquad (16.5)$$

大多数情况下，突变体采取鹰策略的参与者遇见的是采取鸽策略的参与者，得到适应值 V。只有在概率为 ε 下，它遇见的是突变为鹰策略的参与者，得到一个较低的适应值 $(V-C)/2$。对于一个整体都为被赋予鸽策略的群体来说，要想对突变为鹰策略的小部分群体有免疫力，式子（16.4）必须大于式子（16.5），使得采取鸽策略得到的适应值大于采取鹰策略得到的适应值：

$$(1-\varepsilon)\times\left(\frac{V}{2}\right)>(1-\varepsilon)\times V+\varepsilon\times\left(\frac{V-C}{2}\right)$$

这个不等式可以变换为：

$$\varepsilon>\frac{V}{C} \qquad (16.6)$$

不幸的是，当突变体的比例很小（也就是说，当 ε 小于$\frac{V}{C}$时），式子（16.6）就不成立。

如果群体的所有成员都被赋予鸽策略，且有少量的突变体被赋予鹰策略，那些突变体就存活。突变体在很多情况下遇见的是采取鸽策略的参与者，它几乎可以不用开战就获得资源。它遇见另一只突变体的概率很小。大多数情况下，突变体采取鹰策略获得的适应值为 V，而采取鸽策略的参与者获得的适应值为 0。所以，一个都被赋予鸽策略的群体是不稳定的。

总体来说，整体都被赋予为鹰策略的群体是不稳定的，因为它可以被一小部分突变体打败。突变体采取鸽策略避免了在好斗的鹰群体中负伤的危险，但也有全部被赋予鸽策略的群体被一小部分突变为采取鹰策略的突变体打败。面临斗争时，每只采取鸽策略的成员都退缩不战，好斗的鹰就取胜，仅仅靠斗争的意愿就使它们获得了资源。那么，在鹰鸽博弈中，是否存在策略使得群体保持稳定？为了更好地解决这个问题，我们引入了一个新的演化博弈理论核心概念。

16.2 演化稳定策略

迄今为止，我们已经证明整体都为被赋予鹰策略的群体和整体都为被赋予鸽策略的群体都不会处于一个稳定的状态。既然我们把自己限定在所有参与者都采取同样的策略的情况下，那么还有没有其他情况？好吧，让我们回忆一下第 7 章讲的内容，实际上，还有许多可能性，它们都和随机行为有关。让我们思考一种策略，即在概率为 p 时，雄鹿在和其他参与者博弈时采取鹰策略。如果 $0<p<1$，那么有时候它就会退缩（像鸽一样），但其他时候它又站稳立场，随时准备战斗（像鹰一样）。我们试图寻找一个 p 值，在这个值点上群体中的每个参与者都采取鹰策略从而使群体不受侵犯。

因为突变也是随机的，所以不妨思考一下当采取混合策略的参与者遇见另一个采取混合策略的参与者时得到的期望适应值。用 $F(p', p'')$ 来表示被赋予策略 p' 的参与者和被赋予策略 p'' 的参与者博弈时所获得的期望适应值，F 就被定义为：

$$F(p', p'')=p' \times p'' \times \left(\frac{V-C}{2}\right)+p' \times (1-p'') \times V$$
$$+(1-p') \times p'' \times 0+(1-p') \times (1-p'') \times \left(\frac{V}{2}\right) \quad (16.7)$$

为了解释这个表达式，首先注意两个参与者都采取鹰策略的概率为 $p' \times p''$，因为一个被赋予策略 p' 的参与者选择采取鹰策略的概率为 p'，另一个被赋予策略 p'' 的参与者选择采取鹰策略的概率为 p''。所以在概率 $p' \times p''$ 下，每个参与者得到的适应值都为（$V-C$)/2。这就向我们解释了表达式（16.7）中四项的第一项。然后，在概率 $p' \times (1-p'')$ 下，p' 策略类型的参与者采取鹰策略，p'' 策略类型的参与者采取鸽策略，那么采取鹰策略的参与者得到的适应值为 V；这就向我们解释了第二项。在概率为 $(1-p') \times p''$ 时，角色反转，p' 策略类型的参与者获得的适应值为 0。最后，双方都采取鸽策略的概率为 $(1-p') \times (1-p'')$，每个参与者得到的适应值为 $V/2$。

理解了（16.7）这个复杂的方程式（你已经理解了，对吧？），接下来思考在群体中所有成员采取鹰策略的概率为 p，和一部分突变体（该规模为 ε）采取鹰策略的概率为 q，那么采取策略 p 得到的期望适应值为：

$$(1-\varepsilon) \times F(p, p)+ \varepsilon \times F(p, q) \quad (16.8)$$

这只鹿遇见和它采取同策略的概率为 $1-\varepsilon$，遇见突变体采取 q 策略的概率为 ε。采取策略 q 的得到的期望适应值为：

$$(1-\varepsilon) \times F(q, p)+ \varepsilon \times F(q, q) \quad (16.9)$$

当 ε 很小时，要使只包含有被赋予 p 策略的群体能够抵御这种突变，表达式（16.8）的值必须大于表达式（16.9）的值：

$$(1-\varepsilon) \times F(p, p)+ \varepsilon \times F(p, q)>(1-\varepsilon) \times F(q, p)+ \varepsilon \times F(q, q) \quad (16.10)$$

当我们使 ε 无限趋于 0 时，也就是说突变体所占的比例很小，看看表达式（16.10）会发生什么变化。只看表达式的左边部分，我们得出 $(1-\varepsilon) \times F(p, p)$ 的值越来越接近 $F(p, p)$，$\varepsilon \times F(p, q)$ 的值越来越接近于 0。表达式右边部分同样如此。所以，当 ε 越来越接近 0，式子（16.10）表达式就越可以等同于表达式：

$$F(p, p)>F(q, p) \quad (16.11)$$

当 ε 足够接近 0，如果表达式（16.11）成立，那么（16.10）也成立。所以，如果 $F(p, p)>F(q, p)$ 那么在一小部分突变体采取策略 q 的情况下，采取 p 策略的群体是稳定的。现在我们已经进步了！相反，如果 $F(p, p)<F(q, p)$，那么采取策略 p 的群体就很容易被突变体扰乱。

现在利用直觉来分析这些数学表达式可能会有帮助。如果 $F(p, p)>F(q, p)$，那么采取 p 策略的参与者与同类博弈获得的适应值，高于突变体采取 q 策略和采取 p 策略的参与者博弈得到的适应值。因为突变体在数量上很小，采取策略 p 或者 q 的参与者遇见采取 p 策略的可能性，就远远大于遇见采取突变策略 q 的可能性。所以，如果采取 p 策略的参与者和采取 p 策略的参与者博弈得到的适应值大于采取突变策略 q 和采取 p 策略博弈得到的适应值，那么突变体就被淘汰。而如果 $F(p, p)<F(q, p)$ 为真，那么突变体就存活下来。

那么 $F(p, p)=F(q, p)$ 的情况如何呢？就是说采取 p 策略的参与者和采取 p 策略的参与者博弈得到的适应值等于采取突变策略 q 和采取 p 策略博弈得到的适应值。这样的话，不等式（16.10）就变为：

$$\varepsilon \times F(p, q)> \varepsilon \times F(q, q) \text{ 或者 } F(p, q)> F(q, q)$$

观察不等式（16.10），我们发现当 $F(p, p)=F(q, p)$ 时，可以同时减去不等式左右两边的第一项。所以，如果采取 p 策略的参与者和采取 p 策略的参与者博弈得到的适应值等于采取突变策略 q 和采取 p 策略博弈得到的适应值，那么，要使采取 p 策略的参与者对突变体有免疫力，当遇见突变体时，采取 p 的策略得到的适应值就必须大于采取 q 策略得到的适应值。尽管遇见突变体的概率很小，它们决定了采取 p 策略和采取 q 策略可以获得的相对适应值。

总结一下，我们注意到在下列两种情况下，当遇见突变体采取 q 策略时，采取 p 策略都可以使群体保持稳定：（1）当遇见的参与者采取 p 策略时，采取 p 策略得到的适应值高于采取 q 策略得到的适应值；（2）当遇见的参与者采取 p 策略时，采取 p 策略得到的适应值等于采取 q 策略得到的适应值；当遇见的参与者是采取 q 策略时，采取 p 策略得到的适应值高于采取 q 策略得到的适应值。满足这些条件的策略就被称为演化稳定策略。

定义 16.1

演化稳定策略（Evolutionarily Stable Strategy）：在所有的 $q \neq p$ 来说，如果 $F(p, p)>F(q, p)$（强 ESS），或者 $F(p, p)=F(q, p)$ 且 $F(p, q)>F(q, q)$（弱 ESS），那么策略 p 就是演化稳定策略（ESS）。

如果不满足上述任何条件（不论是强 ESS 还是弱

ESS），那么这个策略就不是演化稳定策略。比如，如果 $F(p,p)<F(q,p)$，那么采取突变策略 q 得到的适应值就大于采取 p 策略得到的适应值，采取 q 策略的比例就会呈增长趋势。再者，如果 $F(p,p)=F(q,p)$ 且 $F(p,q)<F(q,q)$，那么采取 q 策略的也会呈增长趋势。尽管采取 p 策略和采取 q 策略获得的适应值相同，当和采取 q 策略的突变体博弈时，采取 q 策略得到的适应值还是大于采取 p 策略得到的适应值。如果 $F(p,p)=F(q,p)$ 且 $F(p,q)=F(q,q)$，那么采取突变策略 q 获得的适应值就等于采取策略 q 获得的适应值，所以在没有外在力量的情况下，采取 q 策略的所占比例既不增大也不减小，即群体处于稳定状态。

我们已经知道了采取什么策略可以使群体保持稳定状态，现在我们就为鹰鸽博弈模型寻找演化稳定策略。利用强 ESS 的条件，我们需要确定是否存在 p 值使得：

$$F(p,p)>F(q,p)\ (q\neq p)$$

利用适应值的显示表达式，这个条件式可以变为：

$$p^2\times\left(\frac{V-C}{2}\right)+p\times(1-p)\times V+(1-p)\times p\times 0+(1-p)^2\times\left(\frac{V}{2}\right)>$$
$$p\times q\times\left(\frac{V-C}{2}\right)+q\times(1-p)\times V+(1-q)\times p\times 0+(1-p)\times(1-q)\times\left(\frac{V}{2}\right),\quad q\neq p \qquad (16.12)$$

为了使这个条件看起来更简单，我们需要对这个条件式做一下调整（这个很简单的），我们就会发现式子（16.12）就等同于：

$$(p-q)\times\left[p\times\left(\frac{V-C}{2}\right)+(1-p)\times V-(1-p)\times\left(\frac{V}{2}\right)\right]>0,\quad q\neq p \qquad (16.13)$$

现在，假定方括号中的项为正值。那么当 $p<q$ 时，条件式（16.13）不成立，因为那样的话 $p-q<0$，所以不等式的左边为负值。如果方括号中的值为负值，同理可证，当 $p>q$ 时，条件式（16.13）也不成立。这样对所有的 $q\neq p$ 来说，条件式都不成立。因而我们得出结论不存在一个 p 值符合强演化稳定策略。

那么让我们再思考一下弱演化稳定策略的条件。首先：

$$F(p,p)=F(q,p),\quad q\neq p$$

通过分析表达式（16.13），当所有的 $q\neq p$ 时，要使 $F(p,p)=F(q,p)$，表达式（16.13）左边等于 0，方括号的项就必须为 0：

$$p\times\left(\frac{V-C}{2}\right)+(1-p)\times V-(1-p)\times\left(\frac{V}{2}\right)=0 \qquad (16.14)$$

调整表达式（16.14），我们发现 $p=\frac{V}{C}$。但如果是这种情况，采取 p 策略的参与者与采取 p 策略的参与者博弈得到的适应值就等于与采取 p 策略博弈时，采取突变策略 q 得到的适应值。于是，$p=\frac{V}{C}$ 就是演化稳定策略的备选值，因为它满足弱演化稳定策略的第一个条件式。第二个条件式是：

$$F(p,q)>F(q,q),\quad q\neq p \qquad (16.15)$$

替换表达式，我们可以将（16.15）变为：

$$p\times q\times\left(\frac{V-C}{2}\right)+p\times(1-p)\times V+(1-p)\times q\times 0+$$
$$(1-p)\times(1-q)\times\left(\frac{V}{2}\right)>q^2\times\left(\frac{V-C}{2}\right)+q\times(1-q)\times V+$$
$$(1-q)\times q\times 0+(1-q)^2\times\left(\frac{V}{2}\right),\quad q\neq p \qquad (16.16)$$

重新调整式子（16.16），将 p 替换为 $\frac{V}{C}$，我们得到：

$$\left(\frac{V}{C}-q\right)\times\left[q\times\left(\frac{V-C}{2}\right)+(1-q)\times V-(1-q)\times\left(\frac{V}{2}\right)\right]>0,\quad q\neq\frac{V}{C} \qquad (16.17)$$

再将这个表达式进行代数调整，我们发现（16.17）就等于：

$$\left(\frac{1}{2}\right)\times\left(\frac{V}{C}-q\right)^2>0,\quad q\neq\frac{V}{C} \qquad (16.18)$$

因为一个数的平方（不论这个数字是正值还是负值）总是正值，条件式（16.18）成立。我们总结出策略 $\frac{V}{C}$ 是弱演化稳定策略。

刚才我们证明的是，采取突变策略 q 与采取 $\frac{V}{C}$ 策略博弈得到的适应值，和采取 $\frac{V}{C}$ 策略与采取 $\frac{V}{C}$ 策略博弈得到的适应值相同；而当与突变体博弈时，采取 $\frac{V}{C}$ 策略得到的适应值就大于采取突变策略得到的适应值。让我们深入探讨为什么是这样的：如果一个突变体选择鹰策略的概率大于 $\frac{V}{C}$，那么当遇见突变体时，它最好不选择鹰策略，因为这样做意味着受到的伤害越小，所以，$\frac{V}{C}$ 策略就优于更具攻击性的突变体策略；如果一个突变体选择鹰策略的概率小于 $\frac{V}{C}$，那么遇见更被动的突变体，最好选择鹰策略，这样做意味着可以掠夺更多的资源，所以策略 $\frac{V}{C}$ 优于突变体策略。

我们终于为鹰鸽博弈找到了使群体保持稳定状态的策略。**如果群体的所有成员选择鹰策略的概率为$\frac{V}{C}$时，那么群体对一小部分突变体的到来就具有免疫力。**让我们从思维跳到在这个“弱肉强食”世界发生的实际的鹰鸽斗争，然后再给博弈加入其他的诱惑因素。

在交配期，加拿大雄性麝香牛就处于鹰鸽博弈，结果是 5%~10% 的麝香牛在斗争中受到致命的伤害（那就是说，有时候采取的是鹰策略）。在有榕小蜂的时候，在雄性麝香牛斗争中有一半的死亡率。而普通蟾蜍的情况就很令人诧异。在交配季节，雄性蟾蜍就在池塘边等雌性蟾蜍来产卵。一旦发现雌性蟾蜍，雄性蟾蜍立马抱住雌性蟾蜍的背。当另一只雄性蟾蜍也发现她时，也会跳到她的背上和第一只雄性蟾蜍撕打。实际上，超过 20% 的雌性蟾蜍身上都会有六只雄性蟾蜍在撕打！有时候，这甚至会导致雌性蟾蜍淹死在池塘里。

16.1 理解测验

当 $V>C$ 时，为鹰鸽博弈模型找一个演化稳定策略。

案例研究

在牛粪上“生存”

在一个星期五的晚上，几个家伙去迪斯科舞厅跳舞。跳了几曲之后，他们发现这里的女的都太瘦。每个人都喝着饮料，思考着是继续待在这里还是重新找个地方。为了引起生物学家的兴趣，我们需要将那些 20 多岁的现代人替换为蜣螂，将迪斯科这个场所替换为牛粪。因为雌蜣螂喜欢在热牛粪上产卵，而雄蜣螂喜欢在牛粪上飞来飞去等待着雌蜣螂的到来。于是，雄蜣螂就必须决定哪片牛粪是它已经夺取的，然后飞到更新鲜的牛粪上。（这种情况类似于找到一个时髦的迪斯科舞厅。）雄蜣螂其次需要考虑的是在什么时候会有一大群雄蜣螂在牛粪上出没。没多久，如果雌蜣螂出现，雄蜣螂必须和其他雄性蜣螂竞争来使卵受精。如果其他雄蜣螂固执地待在原地竞争，那么最好的办法是找另一个没有那么多雄蜣螂出没的牛粪。然而，如果其他雄蜣螂表现出不耐烦的样子，那么它就在原地等待，因为不久那儿就只有它一只蜣螂了。

牛粪上的这几只雄蜣螂在这种策略下的争斗就属于消耗战。每只都在决定是等待（寄希望其他雄蜣螂离开）还是离开。简单来说，假设在一片牛粪上只有两只雄蜣螂，每只都可以等待一分钟或两分钟。（蜣螂就像两岁的小孩子一样没有耐心。）图 16—2 列出了这种情况下的适应值矩阵。

		蜣螂 2	
		一分钟	两分钟
蜣螂 1	一分钟	2，2	2，5
	两分钟	5，2	1，1

图 16—2　蜣螂在牛粪上等待的时间

如果另一只蜣螂待一分钟，那么这只蜣螂就宁愿多待一会获得 5 个适应值。通过打败对手，它成为这片牛粪上的唯一一只雄蜣螂，当雌蜣螂一出现，它就是获胜者。相反，如果另一只蜣螂也待两分钟，那么这只雄蜣螂待两分钟获得的适应值为 1，这样，它待一分钟得到的适应值为 2 就成为比较好的选择。

思考一下备用策略 p，在和同样也采取备用策略的参与者博弈时它获得的适应值为：

$$F(p,p)=p\times2+(1-p)\times[p\times5+(1-p)\times1]$$

在概率为 p 的时候，雄蜣螂停留一分钟获得 2 个适应值（姑且不论其他雄蜣螂停留多长时间）。雄蜣螂停留两分钟的概率为 $1-p$。在这种情况下，如果其他的雄蜣螂停留一分钟，概率为 p，第一只雄蜣螂获得的适应值就为 5；如果其他的雄蜣螂也停留两分钟，概率为 $1-p$，第一只雄蜣螂获得的适应值就只有 1。

要求解一个强演化稳定策略，当都与采取 p 策略的参与者博弈时，采取 p 策略得到的适应值就应该大于采取替代策略 q 得到的适应值：

$$F(p,p)>F(q,p),\ p\neq q$$

替换适应值的显示表达式，得到

$$p\times(1-4p)>q\times(1-4p),\ q\neq p \tag{16.19}$$

很明显，当 p 等于$\frac{1}{4}$时，条件式不成立，因为表达式（16.19）的两边都等于 0。现在思考 p 的值小于$\frac{1}{4}$的情况，那样的话，$1-4p>0$，表达式（16.19）两边都除以 $1-4p$，不等式变为 $p>q$。突变体停留一分钟概率大于 p 时，实际上获得更高的适应值。所以，如果 $p<\frac{1}{4}$，它就不是强演化稳定策略，因为它可以被一些不是很有耐心的

突变体成功侵入。接下来，思考 p 值大于$\frac{1}{4}$的情况，因为 $1-4p<0$，不等式（16.19）的两边除以 $1-4p$，不等式就变为 $p<q$。（记住被除数为负数的时候，不等式的方向要改变。）因此，如果 $p>q$，那么突变体获得更高的适应值。如果 $p>\frac{1}{4}$，它就不是强演化稳定策略，因为它可以被更有耐心的突变体成功侵入。对这个分析做总结，我们注意到蜣螂博弈中不存在强演化稳定策略。

下一步我们转而思考蜣螂博弈中存不存在弱演化稳定策略，满足弱演化稳定策略的第一个条件是当都与采取 p 策略的参与者博弈时，采取 p 策略得到的适应值和采取 q 策略得到的适应值相同：

$$F(p,p)=F(q,p),\ q\neq p$$

继续分析，等式就变为：

$$p(1-4p)=q(1-4p),\ q\neq p$$

当且仅当 $p=\frac{1}{4}$时，条件式成立。满足弱演化稳定策略的第二个条件是，当都与采取突变策略的参与者博弈时，采取 p 策略得到的适应值大于采取突变策略得到的适应值：

$$F(p,q)>F(q,q),\ q\neq p$$

利用显式表达式，我们得到：

$$p\times2+(1-p)\times[q\times5+(1-q)\times1]>$$
$$q\times2+(1-q)\times[q\times5+(1-q)\times1],\ q\neq p \qquad (16.20)$$

简化表达式，用$\frac{1}{4}$代替 p，我们发现（16.20）就等同于：

$$\left(\frac{1}{4}\right)\times(1-4q)^2>0,\ q\neq\frac{1}{4} \qquad (16.21)$$

因为（16.21）一定是真的，所以 $p=\frac{1}{4}$是弱稳定演化策略。

蜣螂选择 25% 的概率停留一分钟的策略是演化稳定策略，因为这样在与采取 p 策略的参与者博弈时，采取 p 策略得到的适应值大于采取突变策略得到的适应值；当与采取突变策略的参与者博弈时，采取 p 策略得到的适应值严格大于采取突变策略得到的适应值。当采取突变策略的参与者比采取 $p=\frac{1}{4}$策略的参与者更没有耐心时，p 策略更优因为它持续的时间比突变体持续的时间长；当采取突变策略的参与者比采取 $p=\frac{1}{4}$策略的参与者更有耐心时，p 策略还是最优策略因为它选择避免一场消耗战。

在这里我想强调一下这个演化稳定策略的特征。蜣螂停留一分钟获得的适应值一定是 2。假定其他蜣螂使用 $p=\frac{1}{4}$策略，停留两分钟获得的适应值还是 2，因为：

$$\left(\frac{1}{4}\right)\times5+\left(\frac{3}{4}\right)\times1=2$$

假定其他雌蜣螂停留一分钟的概率为 25%，对其他任何蜣螂来说，不管停留一分钟还是两分钟获得的适应值都一样。这个事实并非巧合，我们会在下一节解释原因。此刻，我们需要确定不管在迪斯科舞厅（我的意思是牛粪上）停留多久，获得的适应值都是一样的吗？

如果雄蜣螂采取演化稳定策略，那么不管它停留多长时间，得到的期望适应值都是一样的。一项研究估算了交配成功和在牛粪上停留时间长短的关系。结果如图 16—3 所示，横坐标轴表示在牛粪上停留时间的长短，纵坐标轴表示交配成功的情况。结果是很令人惊奇的：不管停留的时间多长，获得的平均适应值都相当稳定，和通过演化稳定策略预测的几乎一致！

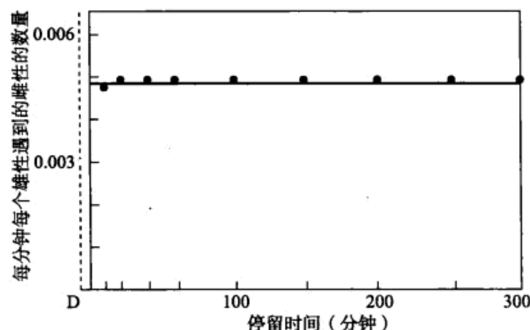


图 16—3　雄蜣螂停留时间长短和交配成功的关系

这项研究的下一步就是戴上一块秒表去迪斯科舞厅，或者干脆我去开一个名为“新鲜牛粪”的迪斯科舞厅。那我简直要发大财了！

16.3 演化稳定策略的特征

如果对侵入有免疫力，我们可从群体利用一种策略来保持稳定，进而发展出演化稳定策略这一概念。正如在本章导读中所暗示的，通过自然选择机制获得最高的适应值，自然“选择”一种策略和理性参与者选择最优策略，从而产生最大的适应值。就这两种选择机制的相似性，我们可以做一个具体的总结。

因为在进化环境中的适应值在社会环境中表现为收益，在二人博弈中 p 策略满足对称纳什均衡的条件为：

$$F(p,p) \geq F(p,q),\ q \neq p$$

换句话说，假定一个参与者采取 p 策略，另一个参与者的最好选择也是采取 p 策略。我们比较这个条件式和策略 p 成为演化稳定策略的条件式：

$$F(p,p) > F(q,p),\ q \neq p$$

或者，

$$F(p,p) = F(q,p),\ q \neq p$$

且：

$$F(p,q) > F(q,q),\ q \neq p$$

所以，如果策略 p 是演化稳定策略，那么或者 $F(p,p) > F(q,p)$，$q \neq p$，在这种情况下，策略 p 为纳什均衡策略；或者 $F(p,p) = F(p,q)$，$q \neq p$，这种情况下，p 策略也是纳什均衡策略。这个结论可以得出以下结果：


点拨

演化稳定策略是纳什均衡。

反过来又如何呢？所有的纳什均衡都是演化稳定策略吗？答案是否定的，演化稳定策略是比纳什均衡更精炼的概念。就是说，**一种策略成为演化稳定策略比成为纳什均衡需要满足更多的条件**。从某种意义上说，演化稳定策略是纳什均衡附加一个稳定条件，这个条件能保证群体在小的冲击下不被侵入的稳定状态。

为了更具体地说明演化稳定策略需要更严格的标准，假设 p 策略满足下列两个条件：

(1) $F(p,p) = F(q,p),\ q \neq p$

(2) $F(p,q) < F(q,q),\ q \neq p$

第一个条件保证了没有比 p 策略更优的策略；所以，采取 p 策略的参与者在博弈中没有改变策略的意识。也就是说，p 策略是纳什均衡。然而，第二个条件表明 p 策略不是演化稳定策略，因为当都与采取 q 策略的参与者博弈时，采取突变策略 q 得到的适应值大于采取 p 策略得到的适应值。（而且，通过第一个条件，当都与采取 p 策略的参与者博弈时，采取 q 策略得到的适应值与采取 p 策略得到的适应值相同）。比如，弱劣策略的对称纳什均衡不是演化稳定策略。要使 p 策略成为演化稳定策略，最少应满足以下条件：不管与采取 p 策略（强演化稳定策略的情况）或者采取 q 策略（弱演化稳定策略的情况）的参与者博弈，采取 p 策略得到的适应值要严格大于采取任何 q 策略得到的适应值，但如果 q 策略弱占优于 p 策略时，那么 p 策略永远不可能优于 q 策略。


点拨

纳什均衡策略不一定是演化稳定策略。特别是当均衡策略在一个对称纳什均衡中是弱劣策略时，这个均衡策略不是演化稳定策略。

严格纳什均衡策略是纳什均衡，在纳什均衡中，每个参与者选择的策略都是最优策略，从而使其他参与者获得一个较低的收益。满足强演化稳定策略的条件和（对称）严格纳什均衡的条件如出一辙。


点拨

一个对称严格纳什均衡是演化稳定策略。

现在我们假设一个演化稳定策略是混合策略，就像在鹰鸽博弈和蜣螂博弈中的情况。回忆第 7 章，如果一个参与者在纳什均衡中随机选择策略，那么她从所有纯策略中随机选择的策略必须使她得到相同的收益；如果觉得所有的选择都无所谓，那么最优的选择就是抛硬币来随机选择。在蜣螂博弈中，25% 概率选择停留一分钟的策略是纳什均衡策略，因为假定其他蜣螂停留一分钟的概率也为 25%，停留一分钟和停留两分钟获得相同的期望收益。在这两种纯策略的任何随机选择获得的期望收益都是 2。现在，既然演化稳定策略一定是纳什均衡策略，如果演化稳定策略是混合策略，那么通过演化稳定策略获得的适应值就等于从纯策略的任何一个选择中获得的适应值。比如，在蜣螂博弈中，如果其他蜣螂采取 $p=\frac{1}{4}$ 策略，采取策略 $\frac{1}{4}$ 得到的适应值就为 2，等于采取其他策略得到的适应值，那就是：

$$F\left(\frac{1}{4},\frac{1}{4}\right)=F\left(q,\frac{1}{4}\right)$$

但这意味着这个演化稳定策略不是一个强演化稳定策略：**采取混合演化稳定策略获得的适应值，等于突变体在纯策略演化稳定策略随机选择策略获得的适应值。**

点拨

如果演化稳定策略不是纯策略，那么它是弱演化稳定策略。

在寻求演化稳定策略集时，这些观点是很有用的。让我们看看这些观点是如何帮助我们分析蜣螂博弈的。因为每一个演化稳定策略是一个对称纳什均衡，不是对称纳什均衡的策略组合不是演化稳定策略。第一步就是寻找所有的对称纳什均衡，因为如果存在任何演化稳定策略，那一定在这些均衡中。在蜣螂博弈中，没有对称纯策略纳什均衡。对称混合策略纳什均衡的定义为：如果另一只蜣螂停留一分钟的概率为 p，那么这只蜣螂就无所谓停留一分钟还是两分钟：

$$2=p\times5+(1-p)\times1$$

调整关于 p 的方程式，得出 $p=\frac{1}{4}$。所以，存在一个特殊的对称纳什均衡，它是 $p=\frac{1}{4}$。这是唯一可能成为演化稳定策略的值。我们知道它不可能是强演化稳定策略，所以我们仅仅需要验证它是否符合弱演化稳定策略的条件。

16.2 理解测验

为图 16—4 的博弈模型找到所有的演化稳定策略。

参与者 2

		慢	快
参与者 1	慢	6，6	0，2
	快	2，0	1，1

图 16—4

案例研究

侧边斑点蜥蜴

为了更进一步探讨演化稳定策略和纳什均衡的关系，让我们返回到 4.2 章节中的石头、布、剪刀博弈。如图 16—5 所示，因为这个博弈被用来解释侧边斑点蜥蜴的颜色，我们用蜥蜴来代替巴特和莉萨，用颜色代替手部运动。我们将对某一时刻的生物环境做进一步的解释。现在，让我们分析这个博弈。

蜥蜴 2

		蓝	橙	黄
蜥蜴 1	蓝	0，0	–1，1	1，–1
	橙	1，–1	0，0	–1，1
	黄	–1，1	1，–1	0，0

图 16—5　侧边斑点蜥蜴

利用第 7 章提到的方法，我们可以知道存在一个特殊的对称纳什均衡，在这个均衡中，每只蜥蜴选择任何一个颜色的概率为 $\frac{1}{3}$。（因为这个有点复杂，我决定省略这个推导过程。）这个就是唯一可以成为演化稳定策略的策略。现在我们想要表明这个纳什均衡不是一个演化稳定策略，就是说在这场博弈中，不存在一个演化稳定策略。

为实现这一目标，思考一下在突变体中，蜥蜴在蓝色和橙色之中随机变色，选择每个颜色的概率为 $\frac{1}{2}$。（不必解释我为什么要选择这个突变，关键是这个突变会存在下来，这已经达到我们的目的了。）令 $F((w, x),(c, d))$ 分别代表当另一只蜥蜴选择混合策略（c, d）时，每只蜥蜴采取混合策略（w, x），即选蓝色的概率为 w，选橙色的概率为 x，选黄色的概率为 $1-w-x$——获得的适应值（同理可证）。

我们想要证明采取策略（$\frac{1}{2}$，$\frac{1}{2}$）可以成功侵入每个成员都采取策略（$\frac{1}{3}$，$\frac{1}{3}$）的群体；换句话说，（$\frac{1}{3}$，$\frac{1}{3}$）不是一个演化稳定策略。为了证明这一点，我们估算当都与采取（$\frac{1}{3}$，$\frac{1}{3}$）策略的参与者博弈时，采取策略（$\frac{1}{3}$，$\frac{1}{3}$）获得的适应值和采取策略（$\frac{1}{2}$，$\frac{1}{2}$）获得的适应值相同。下面是详细的显示表达式：

$$\begin{aligned}F\left(\left(\frac{1}{3},\frac{1}{3}\right),\left(\frac{1}{3},\frac{1}{3}\right)\right)=&\left(\frac{1}{3}\right)\times\left[\left(\frac{1}{3}\right)\times0+\left(\frac{1}{3}\right)\times(-1)+\left(\frac{1}{3}\right)\times1\right]\\&+\left(\frac{1}{3}\right)\times\left[\left(\frac{1}{3}\right)\times1+\left(\frac{1}{3}\right)\times0+\left(\frac{1}{3}\right)\times(-1)\right]\\&+\left(\frac{1}{3}\right)\times\left[\left(\frac{1}{3}\right)\times(-1)+\left(\frac{1}{3}\right)\times1+\left(\frac{1}{3}\right)\times0\right]\\=&\left(\frac{1}{3}\right)\times0+\left(\frac{1}{3}\right)\times0+\left(\frac{1}{3}\right)\times0=0\end{aligned}$$

与：

$$\begin{aligned}F\left(\left(\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{3},\frac{1}{3}\right)\right)=&\left(\frac{1}{2}\right)\times\left[\left(\frac{1}{3}\right)\times0+\left(\frac{1}{3}\right)\times(-1)+\left(\frac{1}{3}\right)\times1\right]\\&+\left(\frac{1}{2}\right)\times\left[\left(\frac{1}{3}\right)\times1+\left(\frac{1}{3}\right)\times0+\left(\frac{1}{3}\right)\times(-1)\right]\end{aligned}$$

$$=\left(\frac{1}{2}\right)\times 0+\left(\frac{1}{2}\right)\times 0=0$$

于是就得到：

$$F\left(\left(\frac{1}{3},\frac{1}{3}\right),\left(\frac{1}{3},\frac{1}{3}\right)\right)=0=F\left(\left(\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{3},\frac{1}{3}\right)\right)$$

所以，策略（$\frac{1}{3}$，$\frac{1}{3}$）要替代突变策略（$\frac{1}{2}$，$\frac{1}{2}$）成为稳定策略，那么当都与采取突变策略的参与者博弈时，采取策略（$\frac{1}{3}$，$\frac{1}{3}$）得到的适应值就必须大于采取策略（$\frac{1}{2}$，$\frac{1}{2}$）得到的适应值，即，

$$F\left(\left(\frac{1}{3},\frac{1}{3}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)>F\left(\left(\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)$$

再次调整表达式，我们得到：

$$\begin{aligned}F\left(\left(\frac{1}{3},\frac{1}{3}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)&=\left(\frac{1}{3}\right)\times\left[\left(\frac{1}{2}\right)\times 0+\left(\frac{1}{2}\right)\times(-1)\right]\\&+\left(\frac{1}{3}\right)\times\left[\left(\frac{1}{2}\right)\times 1+\left(\frac{1}{2}\right)\times 0\right]\\&+\left(\frac{1}{3}\right)\times\left[\left(\frac{1}{2}\right)\times(-1)+\left(\frac{1}{2}\right)\times 1\right]\\&=\left(\frac{1}{3}\right)\times\left(-\frac{1}{2}\right)+\left(\frac{1}{3}\right)\times\left(\frac{1}{2}\right)+\left(\frac{1}{3}\right)\times 0\\&=0\end{aligned}$$

与：

$$\begin{aligned}F\left(\left(\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)&=\left(\frac{1}{2}\right)\times\left[\left(\frac{1}{2}\right)\times 0+\left(\frac{1}{2}\right)\times(-1)\right]\\&+\left(\frac{1}{2}\right)\times\left[\left(\frac{1}{2}\right)\times 1+\left(\frac{1}{2}\right)\times 0\right]\\&=\left(\frac{1}{2}\right)\times\left(-\frac{1}{2}\right)+\left(\frac{1}{2}\right)\times\left(\frac{1}{2}\right)\\&=0\end{aligned}$$

所以，

$$F\left(\left(\frac{1}{3},\frac{1}{3}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)=0=F\left(\left(\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)$$

采取突变策略（$\frac{1}{2}$，$\frac{1}{2}$）和采取策略（$\frac{1}{3}$，$\frac{1}{3}$）获得同样的适应值。

不论是与采取策略（$\frac{1}{3}$，$\frac{1}{3}$）的参与者博弈，还是与采取策略（$\frac{1}{2}$，$\frac{1}{2}$）的参与者博弈，采取策略（$\frac{1}{3}$，$\frac{1}{3}$）得到的适应值和采取突变策略（$\frac{1}{2}$，$\frac{1}{2}$）得到的适应值相同。既然采取策略（$\frac{1}{2}$，$\frac{1}{2}$）的突变体不会被群体淘汰，那么策略（$\frac{1}{3}$，$\frac{1}{3}$）不是一个演化稳定策略。


点拨

不是每场博弈都有一个演化稳定策略。

既然不存在演化稳定策略，群体成员就永远不会固定采取一种策略，因为最终突变体会侵入而且长期存在。纯策略很容易解释这类群体的不稳定性。比如，假设群体所有成员使用纯策略石头。于是，突变体都采取布策略，突变体明显获胜成为主导地位，但它可能被采取剪刀策略的突变者取代。如果采取剪刀策略者获胜，那么突变者使用石头策略可以获胜，如此循环。

在石头、剪刀、布博弈中不存在演化稳定策略的情况，这可以解释侧边斑点蜥蜴循环改变颜色的情况。雄侧边斑点蜥蜴喉咙的颜色不同，这与防御领地的其他特征有关联，因此也与获得的适应值有关联。比如，橙色喉咙的蜥蜴有很强的攻击性，乐于捍卫一片大的领地，而蓝色喉咙的蜥蜴就没那么强的攻击性，它们把自己限定在一片相对小的领地。黄色喉咙的蜥蜴没有自己的领地，相反，它们潜伏在其他蜥蜴的领地里。

现在让我们解释这个生物环境如何和石头、剪刀、布博弈相似。首先，注意更大的领地就意味着有更多的雌性蜥蜴进来，所以附加收益就是更多的繁殖机会。当群体的大多数是橙色喉咙的蜥蜴时，黄色喉咙的蜥蜴好的做法就是聚集在领地的边缘。一旦黄色喉咙蜥蜴的数量足够多，蓝色喉咙蜥蜴就凑过来，通过开发一小片领地来胜过黄色喉咙的蜥蜴。因为蓝色喉咙蜥蜴的领地比橙色喉咙蜥蜴的领地小，相比橙色喉咙蜥蜴来说，它更好地防御了黄色喉咙的蜥蜴。又因为黄色喉咙的蜥蜴没有自己的领地，蓝色喉咙的蜥蜴获得更高的适应值。然而，一旦有大量的蓝色喉咙蜥蜴，橙色喉咙的蜥蜴就能获胜，因为它具更强的攻击性。简单地说，橙色喉咙蜥蜴的扩张性策略使得它败给潜伏在领地边缘的黄色喉咙蜥蜴。相反，黄色喉咙蜥蜴又败给有攻击性的蓝色喉咙蜥蜴，但是不久橙色喉咙蜥蜴表现胜过蓝色喉咙蜥蜴，因为它有更强的攻击性去扩张领地（最初潜伏在周围少量的黄色喉咙蜥蜴会颠覆这一策略）。

通过对侧边斑点蜥蜴群体超过五年的观察，生物学家列出符合这一描述的循环色。循环开始于优势显著的蓝色喉咙蜥蜴，然后是数量急剧扩张的橙色喉咙蜥蜴，最后是数量上增长的黄色喉咙蜥蜴。循环最终回归蓝色喉咙蜥蜴主导权。这一系列事件在没有演化稳定策略的

情况下一直持续，所以群体不会稳定，但是会在不同的基因策略中无限循环。

16.4 多总体模型

在研究过的所有案例中，群体全体成员都处于采取相同策略的环境，不管是侧边斑点蜥蜴为了竞争领地还是雄蜣螂决定在牛粪上停留多长时间。（记住，在进化环境中，“决定”是自然选择机制下“选择”的比喻。）然而，在有些生物环境中，成员被赋予不同的角色。比如，一对雄雌决定该给后代多少关怀。每个成员继承的不仅仅是策略或者与给予关怀度相关的特征，还有其性别赋予它在此环境中的角色。

为了模拟不同角色组成群体的情境，我们假设所有成员是从不同群体选出来的，比如雄性和雌性，在策略组合中，每个群体有一个角色。**演化稳定策略组合（ESSP）是对于群体来说，任何突变体都不适于生存的策略。**

为了更具体地定义演化稳定策略组合，假定有 m 个群体，从而使策略集包含 m 个成员，每个群体中选择一个成员。在雄性群体和雌性群体中选择的话，$m=2$。接下来，定义一个策略组合（$x_1, \cdots, x_m$），这里，x_i 代表群体 i 所采取的策略。F_i（$x_1, \cdots, x_m$）代表群体 i 中的每个成员获得的适应值。现在假设在群体 i 中有一小部分突变体被赋予 y_i 策略。如果（$x_1, \cdots, x_m$）是演化稳定策略，那么采取突变体策略得到的适应值必须小于采取 x_i 策略得到的适应值：

$$F_i(x_1, \ldots, x_m) > F_i(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_m)$$

而且，对 m 个群体中的任何一个成员来说，不等式都成立。

定义 16.2

演化稳定策略组合（Evolutionarily Stable Strategy Profile）：假设有 m 个群体，如果 $F_i(x_1^*, \ldots, x_m^*) > F_i(x_1^*, \ldots, x_{i-1}^*, y_i, x_{i+1}^*, \ldots, x_m^*)$（$y_i \neq x_i^*$，$i=1, 2, 3, \ldots, m$），那么（$x_1^*, \ldots, x_m^*$）是演化稳定策略组合。

注意，这正是严格纳什均衡的定义：**在策略组合中，每个参与者采取的策略都比采取其他策略获得更高的收益。**因为在多总体模型中突变体不和自己博弈，所以在多总体集合下的演化稳定策略组合的定义不同于单群体集合下演化稳定策略的定义。如果在群体 1 中存在突变体，那么突变体获得的适应值仅仅取决于其他 $m-1$ 个群体所采取的策略。更具体一点来说，假设在策略组合方案下有雄性和雌性群体。如果雄性群体部分突变，那么突变体获得的适应值仅仅取决于雌性群体采取的策略；雄性参与者不会遇到其他的雄性（没错的）。所以，突变体不会和另一个突变体博弈。

于是，在多总体集合下的演化稳定策略组合的确定就相对容易了，因为它仅仅意味着找到所有的严格纳什均衡。由于在严格纳什均衡中不存在随机性，所以我们可以只关注纯策略。

案例研究

双亲行为

在每个人类文化的人类学记录中，婚姻是一种规范，家庭是社会组织的核心。所有的父亲都爱他们的子女，这爱比大猩猩的爱多很多，大猩猩似乎都不知道哪些是他们的后代。这种爱使父亲帮助喂养、保护他们的孩子，并教给孩子有用的知识。

双亲行为是一种意在增大他们后代获取的适应值的活动。例如，雄性三棘鱼为鱼卵建巢以抵御周边的危险，这就是双亲行为。大家还知道，雄性把展示卵巢和照看卵当做求爱的一部分，借此向雌性表明他们是多么强大。但这种行为就不是双亲行为，因为它的意图不是增大现存后代的存活率。同样，父亲在晚上给孩子换尿布的行为就属于双亲行为，而把孩子带到公园溜达去约会其他女性就不属于双亲行为。

思考一下动物群体。一个雄性和一个雌性结合，雄性使雌性的卵子受精，自此，战略态势开始。雄雌同时决定是否留下并照看受精卵（借此表现为双亲行为）或者遗弃受精卵以便和其他的异性繁衍后代。更多的受精卵产生和更多的双亲行为意味着更大的繁衍成功率。当有双亲照看行为时，受精卵的存活概率表示为 p_2，p_1 表示只有一方照看时受精卵的存活概率（无论是父亲还是母亲），p_0 表示无双亲照看行为受精卵的存活概率。遗弃受精卵是因为它可能孵化不出后代。获得的适应值就

得益于使更多的卵子受精。假设一个雄性有 r 个机会再次与雌性交配，假设如果遗弃这个受精卵，雌性会产更多的卵，因为她会花更多的时间产卵。更具体地说，如果遗弃受精卵，雌性接下来产 d 个卵子，而如果留下受精卵，仅仅产 s 个卵子，这里 $d>s$。

整合这些我们得到图 16—6 中的适应值矩阵。尽管雄性和雌性有相同的选择集，适应值却因性别不同存在差异。如果双亲都选择留下受精卵，那么雌性产 s 个卵子，每个卵子存活的概率为 p_2，这就是说后代个数的期望值为 $s\times p_2$，这也是一般情况下获得的适应值。如果雄性选择遗弃而雌性留下来照看受精卵，那么她获得的适应值为 $s\times p_1$，因为她产 s 个卵子，每个卵子的存活概率为 p_1，而雄性获得的适应值为 $s\times p_1\times(1+r)$，因为他要提高他的交配率，同理推导其他情况下的适应值。

		雌性：留下来	雌性：抛弃
雄性	留下来	sp_2，sp_2	dp_1，dp_1
	抛弃	$sp_1(1+r)$，sp_1	$dp_0(1+r)$，dp_0

图 16—6　双亲行为

一个演化稳定策略组合是一个严格纳什均衡，组成严格纳什均衡的条件取决于这六个参数的值：d，s，p_0，p_1，p_2 和 r。让我们思考使这四个可能的策略组合中的每一个成为演化稳定策略的条件。

■ **双亲都放弃受精卵。**

要得到演化稳定策略组合，必须满足 $dp_0(1+r)>dp_1$，且 $dp_0>sp_1$。

第一个条件保证了雄性遗弃是更好的策略，第二个条件意味着雌性遗弃后获得更高的适应值。雌性获得的适应值可以表示为不等式 $\frac{d}{s}>\frac{p_1}{p_0}$，这说明通过遗弃受精卵雌性的产卵所得（用 d 和 s 的关系来表示），必须大于有照看行为存在时的存活卵子（用 p_1 和 p_0 的关系来表示）。执行遗弃策略的物种的例子有鲨鱼、鳐形目和电鳐目。另外，照看爬行动物的卵子很费劲。

■ **雌性选择照看，雄性选择放弃。**

当 $sp_1(1+r)>sp_2$，且 $sp_1>dp_0$ 时，这个策略组合是一个演化稳定策略。第一个条件式是关于雄性的。重新调整不等式，得到 $sp_1r>s(p_2-p_1)$。不等式的左边是通过遗弃得到的额外受精卵的期望数量值。记住，r 是当雄性不留下来帮助照看受精卵时他交配机会的增加值。不等式右边是遗弃的成本：如果雄性选择留下来，存活下来的受精卵增加了期望数值。如果遗弃的收益大于遗弃的成本，那么对雄性来说，演化稳定策略就是采取“一夜情”策略。当然，对雄性的计算是在雄性选择留下的基础之上，正如条件式 $sp_1>dp_0$ 反映的那样，如果双亲行为有足够大的影响力，雌性会选择留下照看。这个演化稳定策略组合（雄性放弃而雌性选择留下来）是现在超过 95% 的哺乳动物的选择。

■ **雌性遗弃而雄性选择留下来。**

这是单亲行为的另一种策略组合，当 $dp_1>dp_0(1+r)$，且 $dp_1>sp_2$ 时，它是演化稳定策略组合。可能会发生这种情况，比如当雄性通过遗弃没有那么多机会再繁殖（也就是说，r 的值很小），而雌性通过把注意力都放到产卵上可以产更多的卵子，而非把时间花费在照看卵子上（也就是说，$d>s$）。尽管这种策略组合在我们人类身上不容易发生，在鱼类中却处处可见。大多数鱼类的特点就是单亲照顾，而且常常是雄性照看受精卵。例如，雄性把受精卵含在嘴里来保护它们，这种行为被称为“口育”。

■ **双亲都留下来。**

当 $sp_2>sp_1(1+r)$，且 $sp_2>dp_1$ 时，双亲都照看的行为是一个演化稳定策略组合。如果双亲都留下来对受精卵的存活有很大的帮助，减少交配机会的成本也不大。那么对双亲都留下来来说，这就是演化稳定策略。大约 90% 的鸟类选择双方都留下来孵化受精卵，喂养雏鸡等。在无脊椎动物中，很少有双亲照看幼雏的行为，尽管少于 5% 的哺乳类采取双亲照看策略，却有 30%～40% 的灵长类动物和食肉动物选择双亲照看幼雏。

16.5 敌意的演化

迄今为止，我们已经思考了多成员的群体且突变体只是他们中的很小一部分的群体演变情况。在这一节，假定群体成员数量很小，一小部分的突变体就意味着单独个体采取新的策略。在这类情况下，我们先前的结论就会改变，且更有意思的是，进化可以产生敌意，这是一种伤害别人甚至伤害自己的行为。

返回到单群体的情况下，现在假定有 $n \geqslant 2$ 个成员。用 $F(x,y)$ 表示在使用策略 x 和使用策略 y 博弈时，使用策略 x 获得的适应值。在这个场景中，让我们确定策略 x 成为演化稳定策略的条件。如果突变体的策略为 y，采取策略 x 的参与者获得的适应值为：

$$\left(\frac{n-2}{n-1}\right)F(x,x)+\left(\frac{1}{n-1}\right)F(x,y) \qquad (16.22)$$

采取策略 x 的参与者遇见采取策略 x 的参与者的概率为 $(n-2)/(n-1)$，获得的适应值为 $F(x,x)$；遇见采取策略 y 的参与者的概率为 $1/(n-1)$，获得适应值 $F(x,y)$。

使策略 x 成为演化稳定策略，它获得的适应值就必须大于突变体获得的适应值。注意只有一个突变体，那么突变体遇见的肯定是采取策略 x 的参与者，所以它获得的适应值表示为 $F(y,x)$。这和前面多个成员采取突变策略的情况不同。当且仅当：

$$\left(\frac{n-2}{n-1}\right)F(x,x)+\left(\frac{1}{n-1}\right)F(x,y)>F(y,x),\ y\neq x \qquad (16.23)$$

时，策略 x 是演化稳定策略。不等式（16.23）左边是 $F(x,x)$ 和 $F(x,y)$ 的加权平均数，必须大于突变体在遇见采取策略 x 的参与者时获得的适应值。

现在假设：

$$F(x,x)<F(y,x) \qquad (16.24)$$

在这种情况下，在遇见采取策略 x 的参与者时，采取策略 y 是更优选择。如果一个参与者是理性的，且她期望其他参与者采取策略 x，她肯定会优先选择策略 y。然而，假如 $F(x,y)$ 足够大，即使不等式（16.24）能成立，也不能满足表达式（16.23）的条件。换句话说，在都采取策略 x 的博弈中，即使策略 x 不是最优选择，但它仍是演化稳定策略！

这儿举个有用的例子，改变一下口味，从我们人类中选择一种情况思考。假设有四个公司，在任何给定的时期内，其中的两个公司竞争同一个客户。每个公司都可以选择强势和温和营销方式。前者的成本更大，但会产生更大的效益。适应值（或收益）矩阵如图 16—7 所示：

		公司 1 强势	公司 1 温和
公司 1	强势	1，1	8，3
	温和	3，8	9，9

图 16—7　营销

在这场博弈中，占优策略为温和策略；那么隐含的条件是强势营销方式的成本一定很高。然而，经过下面的证明，**演化稳定策略竟然是强势营销策略！**

若温和策略成为演化稳定策略，当其中一个公司突变采取强势策略时它必须设法应付下来。可以认为"突变策略"是和强势策略博弈的随机决定。利用表达式（16.23）的条件，当且仅当：

$$F(\text{强势，温和})\geqslant\left(\frac{2}{3}\right)\times F(\text{温和，温和})+\left(\frac{1}{3}\right)\times F(\text{温和，强势})$$

时，温和策略不是演化稳定策略。也可以从图 16—7 得到：

$$8\geqslant\left(\frac{2}{3}\right)\times 9+\left(\frac{1}{3}\right)\times 3(=7)$$

不等式的右边是公司采取温和策略获得的适应值，也是表达式（16.22）的具体表现形式。因为有其他三个潜在竞争对手，这个公司遇见另一个公司采取温和策略的概率为 $\frac{2}{3}$，这种情况下，它获得的适应值为 9，遇见另一个公司采取突变策略（强势策略）的概率为 $\frac{1}{3}$，这样获得的适应值为 3。不等式的左边是采取突变策略获得的适应值。作为唯一一个采取强势策略的公司，与它博弈的其他公司采取的当然都是温和策略，所以它获得的适应值为 8。因为采取强势策略的公司获得的适应值为 8，这大于那些采取温和策略公司所获得的适应值 7，突变策略优于那些采取温和策略的公司。所以，温和策略不是演化稳定策略。

注意，如果其他的公司都采取温和策略，在采取强势策略的公司把策略从温和转为强势时，发现它获得的适应值由 9 降到 8。然而，至关重要的是，对其他公司来说，这个转变使得适应值降低得更多，从 9 降到 3。尽

管绝对来说，公司转变为强势策略收益变小，但相对来说，情况还是好转了。从进化方面看，相对结果更重要。所以，如果一个公司从温和策略转变为强势策略，可以发现其他公司也效仿这个公司，因为采取强势策略的公司会比那些采取温和策略的公司赚得更多。(我们仅仅用效仿，而非自然选择，因为通过选择机制，一个更有利的策略可能会优化现在的群体。)

这就是对个人敌意特征在演化方面的观点：你有时会做一些事情降低其他人的福利或者适应值，即使这也会降低你自己的福利或适应值。所以，只要突变策略对其他未采取突变策略的成员有更大的负面影响，那么尽管这个突变策略对自己得到的适应值也有负面影响，它仍然是一个优势策略。

16.3 理解测验

在图 16-7 的博弈模型中，证明采取强势策略是一个演化稳定策略。

本章总结与练习题

本章总结

在演化博弈理论的框架中，这章证明了怎样调整博弈理论来探索动物和我们人类的行为。策略的选择机制并非出于理性，而是自然选择。这就是演化稳定策略概念的动机所在。通过关注稳定的同类群体，我们知道演化稳定策略是一个单一的策略——**对群体来说，演化稳定策略是使群体对任何一次小的冲击都能保持稳定状态的策略。**

尽管建立在自然选择机制而非理性的前提上，演化稳定策略和纳什均衡之间有极强的关联性。演化稳定策略是一个对称纳什均衡，因为如果一个策略不是最好的，那么侵入者采取的更好的策略会获得更高的适应值打破现有的群体结构。不是每个纳什均衡都是演化稳定策略，而是每个严格纳什均衡都是一个演化稳定策略。所以，**演化稳定策略是比纳什均衡更为精炼的概念。**实际上，它是如此严格以致有些博弈中不需要有演化稳定策略。此类例子有和石头、剪刀、布相类似的侧斑点蜥蜴之间的竞争。尽管这些博弈存在纳什均衡，但它不存在一个演化稳定策略。

在自然界中，演化稳定策略被用于多种情景中。在鹰鸽博弈中，从群体中选出的动物为争夺资源而竞争。在蜣螂场景中，在期待雌性蜣螂到来时，每只雄性蜣螂都被赋予一种策略，这种策略决定了蜣螂在牛粪上停留时间的长短。这些都是单个群体的例子，在单个群体中，所有参与者都在群体中相互斗争。在有些场景中，存在多个群体，比如一群雄性和一群雌性相互作用。雄性群体采取的策略集可能不同于雌性群体采取的策略集，或者特定策略获得的适应值的隐含条件因性别不同而不同。**对多总体模型来说，演化稳定策略组合一定是严格纳什均衡。**这个等价关系存在是因为在多总体模型中，突变者一定不会遇见突变者自身。

尽管这章大部分都假定群体规模很大，突变者只占群体的一小部分，但我们也探索了当群体规模很小，很“小”部分的突变者仅涉及群体中的一个成员的情况。一个隐含条件是突变者不会遇见突变者自身，因为只有一个突变者。在这种特殊环境下，演化稳定策略不必是纳什均衡。实际上，一种策略可以是严格劣势策略，但仍然是演化稳定策略！**这类分析强调的是自然选择如何筛选更好的相对表现，而不是更好的绝对表现。**所以，一种也降低自身适应值的策略仍然可以获胜，如果它对其他成员来说意味着更多的损失，因为自然选择机制选择那些相对更适合的策略。

PDG

练习题

1. 思考图 16—8 鹰鸽博弈模型的突变体。与原博弈不同的是鹰伤害鸽的概率为 x，这里 $0<x<1$。求所有的演化稳定策略。（注意答案是关于 x 的条件式。）

赤鹿 2

		鹰策略	鸽策略
赤鹿 1	鹰策略	$\frac{V-C}{2}$，$\frac{V-C}{2}$	V，$-xC$
	鸽策略	$-xC$，V	$\frac{V}{2}$，$\frac{V}{2}$

图 16—8　修改过的鹰鸽博弈

2. 思考图 16—9 中所示的合作博弈，求所有的演化稳定策略。

参与者 2

		a	b
参与者 1	a	1，1	0，0
	b	0，0	2，2

图 16—9

3. 思考图 16—10 中所描述的博弈，求所有的演化稳定策略。

参与者 2

		x	y
参与者 1	x	4，4	2，5
	y	5，2	1，1

图 16—10

4. 思考二人使用两种纯策略的对称博弈，$F(x,y)$ 表示使用混合策略 x 的参与者与使用混合策略 y 的参与者博弈时所获得的适应值。

a. 当 $q\neq p$ 时，用文字表达 $F(p,p)>F(q,p)$ 意味着什么？

b. 当 $q\neq p$ 时，用文字表达 $F(p,p)>F(q,q)$ 意味着什么？

c. 假设 $q\neq p$，$F(p,p)=F(q,p)$ 且 $F(p,p)>F(q,q)$ 时，p 策略是演化稳定策略吗？

5. 思考图 16—11 中所描述的不同性别之间的博弈，求所有的演化稳定策略。

参与者 2

		低	高
参与者 1	低	2，2	3，5
	高	5，3	2，2

图 16—11　不同性别之间的博弈

6. 返回到图 16—2 中的蜣螂博弈。现在假设每只雄性蜣螂停留的时间可以是一分钟，两分钟，三分钟。假定第三分钟时牛粪变得陈腐，这样雌性蜣螂不会飞过来停留。再假设当两只雄性蜣螂停留一分钟或两分钟获得的适应值和图 16—2 中的相同，那么在图 16—12 中，用 w，x，y 和 z 表示适应值的合理值分别是多少？

蜣螂 2

		一分钟	两分钟	三分钟
蜣螂 1	一分钟	2，2	2，5	2，x
	两分钟	5，2	1，1	w，y
	三分钟	x，2	y，w	z，z

图 16—12　蜣螂博弈模型的变体

7. 思考图 16—13 中所描述的博弈，求所有的演化稳定策略。

参与者 2

		w	x	y
参与者 1	w	3，3	2，5	2，2
	x	5，2	1，1	0，0
	y	2，2	0，0	1，1

图 16—13

8. 思考图 16—14 中所描述的博弈，求所有的演化稳定策略。

参与者 2

		x	y	z
参与者 1	x	2，2	1，4	4，6
	y	4，1	2，2	5，2
	z	6，4	2，5	3，3

图 16—14

9. 思考图 16—15 中所描述的博弈，求所有的演化稳定策略。

参与者 2

		w	x	y	z
参与者 1	w	1，1	3，4	0，0	2，3
	x	4，3	5，5	2，4	3，0
	y	0，0	4，2	1，1	0，1
	z	3，2	0，3	1，0	2，2

图 16—15

10. 思考图 16—16 中的鹰鸽博弈。

a. 证明在采取鹰策略概率为 $\frac{2}{3}$ 时是演化稳定策略。

b. 在这章，我们已经规定突变策略是一种单策略。从这个假设出发，思考小部分的突变，其中一半的突变体采取鹰策略，一半突变体采取鸽策略（就是说，他们只使用纯策略）。确定 a 部分的演化稳定策略是否对突变具有免疫力；换句话说，确定这个演化稳定策略是否比这些突变策略获得更高的适应值。

参与者 2

参与者 1		鹰	鸽
	鹰	−1，−1	4，0
	鸽	0，4	2，2

图 16—16 鹰鸽博弈

11. 思考图 16—17 中的不同性别之间的多总体博弈。求所有的演化稳定策略。

雌性

雄性		a	b	c	d
	w	2，0	0，1	3，1	4，2
	x	3，4	6，2	3，3	2，1
	y	1，0	4，3	2，1	2，3
	z	1，2	8，4	4，5	1，3

图 16—17

12. 思考图 16—18 中的不同性别之间的多总体博弈。求所有的演化稳定策略。

雌性

雄性		a	b	c
	w	3，3	1，2	0，3
	x	0，2	4，1	2，1
	y	2，2	6，3	3，1
	z	1，2	3，3	4，0

图 16—18

PDG

第 17 章

演化博弈与生物学：复制者动态模型

学习目标

17.1 复制者动态和鹰鸽博弈理论

17.2 复制者动态的一般定义

17.3 演化稳定策略和复制者动态的吸引子

本章导读

> 以纪念演化论之父达尔文的名义，达尔文奖纪念那些通过把自己的基因从人类基因库中去除，而把福祉留给社会的人。
>
> 设想在一个大学宿舍里堆着脏衣服、性感海报、食品包装袋、课本。在这个杂乱的空间，一个学生随着大声的音乐摇摆。这是一个特别的学生度过的特别的一天。然而据验尸官的报告，当这位特别的学生随着吉它摇摆，在床上蹦跳时，他还想要“做更过分的事情”。李晓蒙，新加坡华商学院的一位学生，因太富于激情以致从三楼的窗户上跳了下去。

尽管在研究本章之前读一下第 16 章的全部内容是有用的，然而也并非必须。关键是阅读 16 章的导读，因为它给我们提供了演化博弈理论概况。如果你已经阅读了这小节，我们就可以继续学习这章内容。

研究博弈论最普遍的方法，也是本书所提倡的，是通过强加条件寻求一种策略组合达到稳定组合的条件。这就意味着一旦参与者使用了这些策略，就没有任何智力或天性的因素来打破这种平衡。这就是运用纳什均衡以及前面章节提到过的演化稳定策略得出的战略。但这种方法缺少的是帮助参与者了解如何实际操作才能达到状态的模型：参与者如何根据稳定组合行动。本章提供了一种可应用自然界策略集的动态模型。这种模型建立在自然选择机制理论上，借助那些更优理论来达到更高效率，从而为下一代组合一种更优策略。（读过 16 章的导读，你就会知道群体中的每个成员都被赋予一种策略，就像她本身就拥有金色头发和创作音乐的天赋一样。）

这种特殊的方法在这本书中被称为**复制者动态**。它具体说明了当可以创造比群体平均期望值高的收益时，群体中使用这个策略的比例就升高；

PDG

而当收益比群体平均期望值低的时候，群体中使用这个策略的比例就降低。借助动物争斗模型，17.1 小节介绍了复制者动态；17.2 小节提供了一个更为正式和普遍的定义；17.3 小节讨论了与演化稳定策略相比较（16 章提到的概念），复制者动态下的策略如何能达到最大的效益。如果你没有读 16 章，你就应该跳过 17.3 小节。本章的其他小节把复制者动态应用到几个生物和社会情境中。我们借用复制者动态来探索猎鹿博弈中的合作猎鹿现象，进而研究过去 120 年来高级联赛协会棒球赛中投球手和击球手的左右手习惯问题。最后，在"重复囚徒困境"背景下，我们来证明合作是如何产生的以及合作什么时候达到稳定状态。换句话说，当社会由骗子组成时，什么时候可以达到稳定状态？而当社会由圣人组成时呢？

17.1 复制者动态和鹰鸽博弈理论

自然界的许多动物争斗都与资源掠夺有关，比如领地的占有，或者是与雌性动物的交配机会。为了把如此复杂的场景简单化，生物学家已经建立起鹰鸽博弈理论模型。在这场博弈中，两个动物之间进行博弈，每种都被赋予可能用到的两种策略中的一种。采取鹰策略（强硬不退让）的动物希望通过争斗来获得对资源的控制权。相反，采取鸽策略（妥协退让）的另一只动物也做好了争斗的架势，好像要应战的样子，但是当其他动物来决斗时它又退后放弃。从适应值方面来说，假定资源值为 V。如果一只鹰和鸽子争斗，鹰就获得对资源的控制权，因为一旦鹰开始争斗，鸽子就放弃了。鹰就得到了适应值 V，而鸽子获得的适应值为 0。如果两个潜在的队手都是鸽子式的心态，那么这个争斗的态势一直持续到一方放弃。如果我们假定双方都有 50% 的概率坚持的时间比另一方长，当两方决斗时，期望的适应值就是$\frac{V}{2}$。最后，如果两只鹰决斗，争斗持续直到其中一方受伤。失败者的适应值就是 $-C$（这里的 C 是受伤的成本），而争斗中的胜利者就获得了资源，得到适应值为 V。假定博弈的双方都是采取鹰策略的参与者，它们获胜的概率一样大，那么它们的适应值就是（$V-C$）/2。

假定每只动物接受的适应值基线为 B，$B>0$。那么它们得到的适应值就是 B 加上它们在博弈中得到的适应值。最后，假定 $C>V>0$，那么受伤成本的适应值就大于得到领地利益的适应值，再假设 B 大到可以使得到的适应值都为正，即能满足 $B>(V-C)/2$。这个场景给我们提供的战略形式如图 17—1 所示：

动物 2

动物 1	鹰	鸽
鹰	$B+\frac{V-C}{2}, B+\frac{V-C}{2}$	$B+V, B$
鸽	$B, B+V$	$B+\frac{V}{2}, B+\frac{V}{2}$

图 17—1　基于适应值基线的鹰鸽博弈理论

演化博弈理论研究的重点是群体，而非单个动物。目的是描述稳定群体所采取的混合策略。为了达到这个目的，首要关注的是策略，从这个角度出发，动物仅仅是研究策略的工具。使用策略的动物可以随意挑选，但群体采取的混合策略可以一直存在。

我们用来描述稳定群体结构的方法是以自然选择理论为基础的：采取获得较高适应值的策略可以繁殖更多的后代，这就增加了后代群体采取该策略的概率。把这种机制模型化，我们就需要记录采取每种策略可以得到的适应值。与另一只被赋予 y 策略的动物博弈时，一种被赋予 x 策略的动物得到的适应值用 $F(x, y)$ 表示，即 F（鹰，鸽）$=B+V$。

思考那些被赋予鹰策略的动物：当与采取相同策略的另一个参与者争斗时，每个参与者获得 $B+(V-C)/2$ 的适应值；当与采取鸽策略的参与者争斗时，获得 $B+V$ 的适应值。所以，在这个群体中，采取鹰策略的动物可以得到的期望平均适应值取决于这个群体的混合策略。令 h^t 表示在 t 代中被赋予鹰策略的动物在群体中的比例，采取鹰策略得到的平均适应值就表示为：

$$F^t(\text{鹰})=h^t\times\left(B+\frac{V-C}{2}\right)+(1-h^t)\times(B+V)$$

鹰遇见采取同样策略的动物的概率为 h^t，获得的适应值为 $B+(V-C)/2$。采取鹰策略的动物遇见采取鸽策略动物的概率为 $1-h^t$，采取鹰策略的动物获得的适应值

为 $B+V$。同理，在 t 代中的采取鸽策略得到的适应值表示为：

$$F^t(鸽)=h^t\times B+(1-h^t)\times\left(B+\frac{V}{2}\right)$$

复制者动态解释了当采取策略获得的适应值大于群体平均适应值时，使用这种策略的后代的比例就会增长，反之则减少。在鹰鸽博弈理论中，在后代 t 中群体平均适应值表示为：

$$\bar{F}^t=h^t\times F^t(鹰)+(1-h^t)\times F^t(鸽)。$$

群体中被赋予鹰策略成员比例为 h^t，得到的平均适应值为 F^t（鹰），而群体中被赋予鸽策略的成员比例为 $1-h^t$，得到的平均适应值为 F^t（鸽）。

在鹰鸽博弈理论的框架下，复制者动态方程就表示为：

$$h^{t+1}=h^t\times\left(\frac{F^t(鹰)}{\bar{F}^t}\right) \tag{17.1}$$

这个方程式说明了 $t+1$ 代中鹰的比例，表示为 h^{t+1}，就等于在 t 代中的比例乘以采取鹰策略的适应值与群体平均适应值的比率。如果采取鹰策略得到的适应值大于群体的平均适应值，即 $F^t(鹰)>\bar{F}^t$，那么 $F^t(鹰)/\bar{F}^t>1$，所以根据方程（17.1），$h^{t+1}>h^t$，所以采取鹰策略的参与者在群体中的比例升高了。而且，采取鹰策略得到的适应值比群体平均适应值越大，采取鹰策略的参与者的比例升高越快。相反，如果采取鹰策略得到的适应值小于群体的平均适应值，那么 $F^t(鹰)/\bar{F}^t<1$，又根据等式（17.1），$h^{t+1}<h^t$，采取鹰策略参与者的比例就下降了。

正如复制者动态所解释的那样，群体如何演化取决于群体中被赋予鹰策略的数量和被赋予鸽策略的数量。接下来，我们的任务就是描述这个演化过程。假如在群体中的当前混合策略下，采取鹰策略得到的适应值大于群体的平均适应值：

$$F^t(鹰)>h^t\times F^t(鹰)+(1-h^t)\times F^t(鸽)$$

不等式两边都减去 $h^t\times F^t$（鸽），我们得到不等式：

$$(1-h^t)\times F^t(鹰)>(1-h^t)\times F^t(鸽)$$

两边都除以 $(1-h^t)$，我们得到：

$$F^t(鹰)>F^t(鸽)$$

最后，把你的魔杖指到这一页然后说“统统石化”。（如果你不是巫师，那就省了这一步吧。）当且仅当采取鹰策略得到的适应值大于采取鸽策略得到的适应值时，采取鹰策略得到的适应值大于群体平均适应值。既然只有两种策略，如果鹰的策略优于平均执行策略，那么采取鹰的策略优于采取鸽的策略。

从复制者动态方程看出，当采取鹰策略得到的适应值大于采取鸽策略得到的适应值时，采取鹰策略的成员在群体中的比例升高，它表示为：

$$h^t\times\left(B+\frac{V-C}{2}\right)+(1-h^t)\times(B+V)>h^t\times B+(1-h^t)\times\left(B+\frac{V}{2}\right) \tag{17.2}$$

两边取消 B，得到关于 h^t 的不等式：

$$h^t<\frac{V}{C}$$

所以，当 $h^t<\frac{V}{C}$ 时，群体中采取鹰策略的成员的比例升高。记得我们假定过 $\frac{V}{C}<1$，即受伤成本大于占有领土利益的值。当采取鹰策略的成员比例足够低，具体来说就是 $h^t<\frac{V}{C}$ 时，采取鹰策略得到的适应值大于采取鸽策略得到的适应值，根据复制者动态方程，采取鹰策略的成员在群体中的比例升高。

相反，如果采取鸽策略得到的适应值大于采取鹰策略得到的适应值，那么采取鹰策略得到的适应值就小于群体平均适应值，采取鹰策略的成员的比例就降低。如果反转不等式（17.2），解关于 h^t 的不等式，我们会发现当 $h^t>\frac{V}{C}$ 时，采取鹰策略的成员比例降低了。如果整个群体对采取鹰策略的成员更严厉，那么采取鸽策略得到的适应值就会更高，由复制者动态方程我们知道，采取鸽策略成员的比例会升高。

把所有这些解释放在一起，我们就可以描述出随着时间推移群体的演化过程：

如果 $h^t<\frac{V}{C}$，那么 $h^{t+1}>h^t$。

如果 $h^t=\frac{V}{C}$，那么 $h^{t+1}=h^t$。

如果 $h^t>\frac{V}{C}$，那么 $h^{t+1}<h^t$。

这组条件式可以描绘成图 17—2，当 $h^t<\frac{V}{C}$，h^t 就增大，且接近 $\frac{V}{C}$，如图中的箭头向右所指向 $\frac{V}{C}$。如果 $h^t>\frac{V}{C}$，那么 h^t 就减小，如我们看到的箭头向左指向 $\frac{V}{C}$。于是，复制者动态把采取鹰策略的比例推向 $\frac{V}{C}$。当 $h^t=\frac{V}{C}$，那么 $h^{t+1}=h^t$，采取鹰策略的成员比例不变。尽管你认为可能是因为用了魔杖施了魔法，但实际上是因为采取鹰策略得到的适应值正好等于采取鸽策略得到的适应值，也恰好等于群体的平均适应值。既然鹰在大体上既不处于

优势也不处于劣势，那么它的比例既不上升也不下降。

在群体结构中，采取鹰策略成员的比例等于$\frac{V}{C}$就是一个静止点。静止点就是使得复制者动态中不受干扰的群体混合策略。一旦群体处于静止点，复制者动态就使群体一直处于这个状态，一代又一代。①

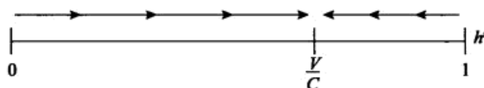


图 17—2　鹰鸽博弈理论中的复制者动态

> **定义 17.1**
>
> **静止点（rest point）：如果群体结构一旦处于点 S^*，且群体结构一直稳定在这个点，那么群体结构 S^* 就是一个静止点。**

群体结构一开始可能处于静止点，但复制者动态不一定就能导致群体达到这个点。若要发生这种情况，静止点必须有局部稳定性。局部稳定静止点（或者说吸引子）是一个静止点，如果群体结构一开始接近这个静止点，它最终会停留在这个静止点。只有当群体一开始接近这个静止点，群体结构才能收敛到这个静止点，所以说它是局部稳定。

> **定义 17.2**
>
> **吸引子（attractor）：如果当群体结构一开始接近点 S^*，且最终群体结构会走向点 S^*，那么这个静止点 S^* 就是一个吸引子。**

需要强调的是，静止点和吸引子都是相对动态过程所描述的特殊点。改变了动态演化过程，那些为局部吸引子的群体结构也随之发生改变。这一章我们将着重关注静止点和复制者动态下的吸引子。

为了更好地理解隐含在静止点和吸引子中的稳定性，思考图 17—3。

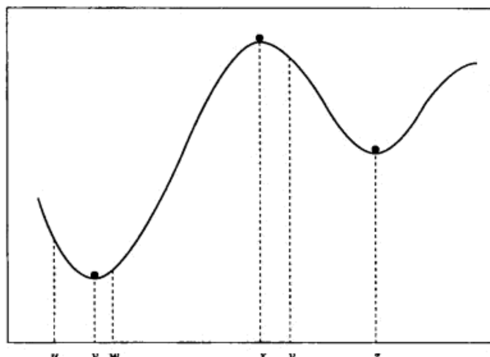


图 17—3　静止点和吸引子的稳定性

设想一个球沿着一个维度的平面运动，在这个平面，动态参数是重力，而非自然选择。静止点就是在一个位置，球再也不运动。位置 u 不是静止点，因为如果球处在 u 的位置，它还会朝着 v 的方向向下滚动。现在假定球在 v 位置，再也没有任何力使它远离 v 位置；一旦滚到那儿，它就停在那儿。位置 v 点就是一个静止点。另一个静止点是位置 x，一旦球被置于山的顶端，它就一直停留在那。然而，v 和 x 都是静止点，但 v 是吸引子而 x 不是。注意我们如果把球拨动一点点，使它偏离 v（比如朝着 w 的方向拨动）球还是会朝着 v 的方向向下滚动并最终停留在那，所以 v 是一个吸引子。相反，如果我们在位置 x 把球朝着 y 的方向拨动，球会向下滚动，偏离 x 滚向 z。即使球从接近 x 的位置开始滚动，最终也不会停留在 x。所以说尽管 x 是一个静止点，但不是吸引子。

17.1 理解测验

为图17—4中的博弈找到吸引子。

参与者 2

参与者 1	跳	跑
跳	3,3	2,1
跑	1,2	2,2

图 17—4

---

① 详细来说，根据复制者动态，全部采取鸽策略的群体结构也是一个静止点，因为如果$h^t=0$，那么$h^{t+1}=h^t\times[F^t（鹰）/\bar{F}^t]=0\times[F^t（鹰）/\bar{F}^t]=0$，所以，$h^{t+1}=0$。换句话说，如果群体中没有采取鹰策略的成员，那么采取鹰策略就不能繁衍后代。类似的，全部要取鹰策略的群体结构也是一个静止点，因为$h^t=0$，则$h^{t+1}=h^t\times[F^t（鹰）/\bar{F}^t]=1\times1=1$。——作者注

被赋予采取鹰策略的成员所占比列为$\frac{V}{C}$的群体结构就是一个吸引子。实际上，它满足了更严格的特征：不管群体结构一开始的组成比例是什么（只要群体成员是由鹰和鸽组成），这个群体就最终会演化为采取鹰策略的成员比例为$\frac{V}{C}$，甚至都不需要将开始时采取鹰策略的比例就定位在接近$\frac{V}{C}$。如图 17—2 所示，群体趋向于$\frac{V}{C}$就如图 17—3 中球趋向于向谷底滚动。

用动物争斗作为模型，借助鹰鸽博弈理论，复制者动态预测群体由混合类型成员组成。有的有攻击性，即使负伤也要战斗，有的成员摆出战斗的姿势，却会临时退却。当占有领土的利益越大（即 V 越大）时，攻击性类型的比例$\frac{V}{C}$就越高；而当负伤成本越大（即 C 越大）时，攻击性类型的比例$\frac{V}{C}$就越低。它解释了当情况越危险而负伤成本越小时，群体更具攻击性。

17.2 复制者动态的一般定义

思考一场博弈中有 m 种策略，令 $F_i(r_1^t,\cdots\cdots,r_m^t)$ 表示当 t 代中群体结构为 $(r_1^t,\cdots\cdots,r_m^t)$ 时，采取策略 i 得到的平均适应值；那就是说，群体中被赋予策略 1 的比例为 r_1^t，被赋予策略 2 的比例为 $r_2^t,\cdots$，被赋予策略 m 的比例为 r_m^t。记住比例和为 1：即 $r_1^t+r_2^t+\cdots\cdots+r_m^t=1$。假定适应值总是正数：

对所有的 $(r_1^t,\cdots,r_m^t)$ 来说，$F_i(r_1^t,\cdots,r_m^t)>0$

在所有的博弈中，只要将适应值基线定的足够大，这个不等式都能成立。最后，把群体得到的平均适应值定义为 $\overline{F}_i$，或者：

$$\overline{F}(r_1^t,\cdots,r_m^t)=r_1^t\times F_1^t+r_2^t\times F_2^t+\cdots+r_m^t\times F_m^t$$

这里：

$$F_i^t=F_i(r_1^t,\cdots,r_m^t)$$

比如在鹰鸽博弈理论中，如果鹰采取策略 1，鸽采取策略 2，于是：

$$F_1(r_1^t,r_2^t)=r_1^t\times\left(B+\frac{V-C}{2}\right)+r_2^t\times(B+V)$$

根据复制者动态，策略 i 在下一代的比例为：

$$r_i^{t+1}=r_i^t\times\left[\frac{F_i(r_1^t,\cdots,r_m^t)}{\overline{F}(r_1^t,\cdots,r_m^t)}\right]$$

当采取这种策略带来的适应值大于群体平均适应值的时候，被赋予这种策略的成员比例就升高。

$$\frac{F_i(r_1^t,\cdots,r_m^t)}{\overline{F}(r_1^t,\cdots,r_m^t)}>1$$

当采取这种策略带来的适应值等于群体平均适应值的时候，被赋予这种策略的成员比例就保持不变。

17.3 演化稳定策略和复制者动态的吸引子

在 16 章中，已经提出了一种复制者动态的替代方法。演化稳定策略在理性博弈理论上与纳什均衡相似。正如纳什均衡规定一个策略组合是稳定状态（没有任何参与者想要改变现有策略）的条件一样，一个演化稳定策略是这样一个策略：如果群体的所有成员采取这个策略，那么群体结构是稳定的，即任何突变体在侵入群体时都是失败的。更具体地说，**采取突变体策略得到的适应值低于采取当前策略得到的适应值。**

本节的主要任务就是比较演化博弈理论的两种方法，即演化稳定策略和复制者动态的相同和不同点。我们将演化稳定策略的使用限定在由同类组成的群体中，也就是说，群体中的所有成员采取相同的策略。尽管一个突变体会明显采取一个不同的策略，但我们假设所有的突变体都采取相同的突变策略。所以，在以演化稳定策略点上的共同策略为特征的群体中，我们仍然可以观察到多样行为，因为我们允许采取混合策略。比如，在鹰鸽博弈中，如果选择鹰策略的概率为 p 是一个演化稳定策略，那么在任何时候，可以观察到群体中表现出鹰行为的比例为 p，群体中比例为 $1-p$ 的表现出鸽行为。

尽管复制者动态可以允许采取混合策略，但我们将这章的重点放在纯策略上。为了抵消这个限定性的假设，我们允许稳定群体结构的多样性。回到鹰鸽博弈中，从复制者动态下的稳定群体中，我们得出比例为 p 的成员采取鹰策略，比例为 $1-p$ 的成员采取鸽策略。

从某种意义上来说，从群体可以表现的行为来看，演化稳定策略方法和复制者动态方法是一样的。以鹰鸽博弈为例，我们可以发现群体层面上，以下两个条件是等同的：（1）每个成员都采取混合策略，而在混合策略中，采取鹰策略的概率为 p；（2）群体中比例为 p 的成员可以使用纯策略鹰策略。现在，假设一个动物群体赤鹿被置于情境 1，另一个动物群体榕小蜂被置于情境 2。如果你是记录这两个群体行为的一个生物学家，你会发现这些记录很相似：在每个群体中，比例为 p 的成员选

择鹰策略。从这个意义上说，在两种情境下，也就是同种群体使用混合策略和异种群体使用纯策略的情况毫无二致。

这种等同效应还体现在稳定群体上。在前面的小节中，已经证明复制者动态的一个吸引子是群体的$\frac{V}{C}$采取鹰策略。在 16.2 小节中，我们也证明了鹰鸽博弈中独特的演化稳定策略是混合策略，其成员采取鹰策略的概率为$\frac{V}{C}$。这两个演化博弈理论的方法表达了同一个意思：被观察到的博弈中比例为$\frac{V}{C}$的成员采取鹰策略。尽管是从不同群体中得出的策略，结果是一样的。

鹰鸽博弈中的这种等同不是偶然的，因为在复制者动态下，演化稳定策略和吸引子之间有很强的关联性。


点拨

在复制者动态下，每个演化稳定策略都是一个吸引子。

这个命题反过来就不一定真，因为复制者动态的吸引子不需要是演化稳定策略。然而，如果我们把博弈限定在两个参与者和两种策略中，比如鹰鸽博弈，那么演化稳定策略就等同于复制者动态下的吸引子。


点拨

在只有两个参与者且每个参与者能使用两种策略的情况下，每个演化稳定策略都是一个吸引子，每个吸引子都是一个演化稳定策略。

借助复制者动态，让我们来研究多种场合下的情况。猎鹿博弈探索了合作猎鹿的产生。在鹰鸽博弈中只有一个吸引子，而在猎鹿博弈中有一对吸引子。还有一种完全不同的情况：美国职棒大联盟中击球手和投手的左右手习惯的演化。有趣的是，在我们的模拟中不存在吸引子，且预计有循环策略。最后的一个应用是我们经常提到的："重复囚徒困境"。在这个博弈中，要么每个人都得到最好的结果，要么得到最坏的结果。

案例研究

猎鹿博弈

许多物种通过合作达到一个共同的目的。狮子（通常是雌性）群体逐猎时比单个行动更有效。在追逐如羚羊、牛羚时，有的狮子就直接朝着其他狮子等待（等羚羊靠近就捕获）的方向追赶。土狼既单独捕猎也群体捕猎，它们合作规模大小取决于追逐的猎物。通过共同合作，土狼可以有效地捕捉到更大的猎物，如斑马。

刚刚描述的战略情景可以模拟为猎鹿博弈，如图 17—5 所示，在猎鹿博弈中，群体的两个成员单独或合作捕猎。正如 18 世纪哲学家卢梭所描述的，两个猎手既可以合作捕捉更大的猎物，如鹿，又可以单兵作战得到小的猎物，如兔子。从适应值来说，都选择鹿的结果（得到适应值为 4）要好于选择追逐兔子的结果（得到的适应值为 3）。

		参与者 2	
		猎物为鹿	猎物为兔
参与者 1	猎物为鹿	4, 4	1, 3
	猎物为兔	3,1	3,3

图 17—5　猎鹿博弈

假设在 t 代中，群体中比例为 s^t 的成员采取合作的策略。那么每个成员采取合作策略得到的平均适应值为：

$$s^t\times 4+(1-s^t)\times 1=1+3s^t$$

而如果单兵作战的话，得到的平均适应值为：

$$s^t\times 3+(1-s^t)\times 3=3$$

借助复制者动态，当且仅当采取合作策略得到的适应值大于群体的平均适应值时，采取合作策略的成员的比例才会增加。因为只有两种策略，这就是说，采取合作策略得到的适应值大于采取单兵作战策略得到的适应值。换句话说，当且仅当：

$1+3s^t>3$，或者 $s^t>\frac{2}{3}$时，$s^{t+1}<s^t$

或者，当且仅当采取单兵作战策略得到的适应值大于采取合作策略得到的适应值时，采取合作策略的成员的比例才会减少。就是说，当且仅当：

$$3>1+3s^t，或者\ s^t<\frac{2}{3}时，s^{t+1}<s^t$$

这个动态过程如图 17—6 所示：当 $s^t<\frac{2}{3}$时，采取合作策略的成员减少，当 $s^t>\frac{2}{3}$时，采取合作策略的成员增加。借助复制者动态，在一个群体中，所有成员都采取合作策略就是一个吸引子。如果当前群体中采取合作策略成员比例非常小（即如果 $s^t<\frac{2}{3}$），那么这个比例将继续下降，因为采取单兵作战策略得到的适应值大于采取

合作策略得到的适应值；在图 17—6 中，当 $s^t<\frac{2}{3}$时，箭头所指的方向就是采取合作策略成员的比例。相反，如果 $s^t>\frac{2}{3}$，那么采取合作策略的成员增加，而且根据复制者动态，在一个群体中，所有成员采取合作策略是一个吸引子。在图 17—6 中，箭头指向所有成员采取合作策略的方向。

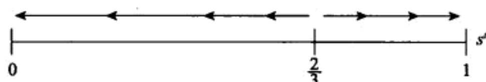


图 17—6　猎鹿博弈中的复制者动态

这些结果很有意义。只有当你的博弈对手也想要得到更多时，你才值得采取得到更多的行动，因为只有这样你的行动才最有效。那么根据复制者动态，只有当一开始采取合作策略的成员足够多，采取合作策略的成员才会一直增加。然而，如果这个比例一开始很小，那么更优策略就是采取单兵作战策略。于是，因为越来越多的成员采取单兵作战的方式，这就使得采取单兵作战方式更有吸引力，因为只有很少的成员会选择合作。在猎鹿博弈中，产生于复制者动态的稳定群体结构取决于刚开始的混合比例。

还有最后一种我们还没有分析的群体结构：群体一开始有$\frac{2}{3}$的成员采取合作的策略，$\frac{1}{3}$的成员采取单兵作战的策略。在这样的群体中，采取合作与采取单兵作战策略得到的适应值都为 3；所以，采取每种策略得到的适应值都和群体平均适应值相同（当然都是 3）。既然所有的策略得到的适应值都一样，复制者动态就使群体处于稳定状态。一个群体有$\frac{2}{3}$的成员采取合作策略就是一个静止点。然而，它不是一个吸引子。为了证明这一点，假设在群体中有些微的变动，比如，一种致命的疾病在群体蔓延，且由于偶然因素，在采取单独作战策略的成员中死亡率要更高些。现在采取合作策略的成员的比例就略高于$\frac{2}{3}$，通过图 17—6，群体结构将稳定发展直到所有的成员都采取合作策略。因为群体结构回不到群体中$\frac{2}{3}$成员采取合作策略的静止点，所以群体刚开始有$\frac{2}{3}$的成员采取合作策略不是一个吸引子。

在猎鹿博弈中，复制者动态预测一个稳定群体的结构要么是所有成员都采取合作策略，要么是所有成员都采取单兵作战策略，而这些都与刚开始的群体结构比例有关，这导致了一个特殊的静止点，也就是吸引域中的那个静止点。

> **定义 17.3**
>
> **吸引域（basin of attraction）：一个静止点 s^* 的吸引域是这样一个域，如果群体结构一开始的比例在这个域中，那么最终的群体结构还在 s^* 点上。**

在图 17—6 中，“所有成员都采取单独作战策略”吸引子的吸引域包括的范围是少于$\frac{2}{3}$的成员采取合作策略，而“所有成员都采取合作策略”吸引子的吸引域包括的范围是多于$\frac{2}{3}$的成员采取合作策略。静止点（群体中$\frac{2}{3}$成员采取合作策略，$\frac{1}{3}$成员采取单兵作战策略）的吸引域是一个分离的点：群体中$\frac{2}{3}$成员采取合作策略，$\frac{1}{3}$成员采取单兵作战策略。因为它的吸引域这么小，能达到这样的静止点的机会确实很渺茫。

17.2 理解测验

为图17—7中的博弈找到复制者动态下的吸引子。

参与者 2

参与者 2		黑	白
	黑	5, 5	1, 1
	白	1, 1	2, 2

图17—7

棒球中的左右手习惯

尽管由生物环境所决定，演化博弈理论也可以对人类社会行为产生影响。所以，让我们把复制者动态应用到探讨棒球运动员的左右手的历史格局中。如前面所讲的，右撇子击球手比左撇子投手表现好，而左撇子击球手比右撇子投手表现好。图 17—8 反映了从 1876 年到 1985 年这 110 年来美国职棒大联盟中右撇子击球手和左撇子投手的比例。

数据反映了一个耐人寻味的历史趋势。开始大约 85%～90% 的击球手和投手是右撇子，这和在普通人群中右撇子的概率一样。然而，慢慢的右撇子击球手和投

手的比例随着时间下降了。另外一个值得注意的是击球手和投手的群体结构似乎朝一个方向变化。结构最后就趋于不变的水平，所以今天大约 65% ~75% 的击球手和投手是右撇子。

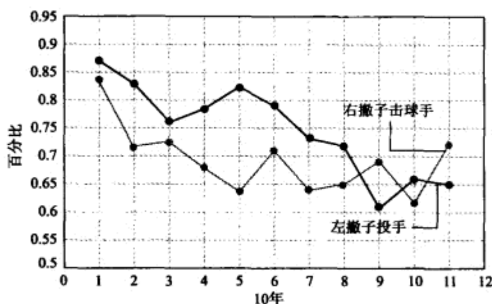


图 17—8　右撇子击球手和左撇子投手，1876-1985 年

让我们看演化博弈理论是否能解释这一趋势。和这章前面的模型不同，现在我们假设根据复制者动态，有两个群体共同演化。一个群体由既能用左手又能用右手击球的击球手组成。另一个群体由既能用左手又能用右手投球的投手组成。图 17—9 列出了一个投手和一个击球手博弈的情况。注意右撇子击球手比左撇子投手表现好，而左撇子击球手比右撇子投手表现好，而这也和事实一致（看表 2—2）。

		投手	
		右撇子	左撇子
击球手	右撇子	30, 30	44, 24
	左撇子	36, 21	30, 30

图 17—9　棒球

我们的目的是理解击球手（从垒的左边还是右边击球）的结构和投手（用右手投或者是用左手投）的结构。这样做的时候，我们还是使用自然选择机制。一个投手或击球手并不能自己选择用左手还是右手，而是被指定用左手还是右手。如果说右撇子击球手比左撇子击球手表现更好，那么教练就会把左撇子击球手换成右撇子击球手，这样右撇子击球手的比例就上升了。

在列出复制者动态方程之前，让我们在这里先用一点直觉。观察图 17—9，我们应该预测存在一个既用左手又用右手类型的结构。如果所有的投手都是右撇子，那么左撇子击球手胜于右撇子击球手，得到的适应值为 36（和 30 比较）。但如果这样导致所有的击球手都是左撇子的话，那么我们应该预测教练开始裁掉右撇子投手换成左撇子投手，因为后者的表现优于前者（比较 30 和 21）。所以，所有的击球手使用手的习惯相同或者所有的投手使用手的习惯相同都不是一个静止点。

为了得到这个场景下的复制者动态，令 b^t 表示 t 代中右撇子击球手的比例，p^t 表示右撇子投手的比例。那么右撇子击球手的平均适应值（或表现）为：

$$F_B^t（右撇子）=p^t\times 30+(1-p^t)\times 44=44-14p^t$$

右撇子击球手遇见右撇子投手的概率为 p^t，他得到的适应值为 30；遇见左撇子投手的概率为 $1-p^t$，且右撇子击球手的适应值明显高，为 44。那么左撇子击球手的平均适应值为：

$$F_B^t（左撇子）=p^t\times 36+(1-p^t)\times 30=30+6p^t$$

那么击球手的复制者动态为：

$$b^{t+1}=b^t\times\left(\frac{44-14p^t}{b^t\times(44-14p^t)+(1-b^t)\times(30+6p^t)}\right) \tag{17.3}$$

这里的分母是所有击球手的平均适应值。因为只有两种策略，所以当右撇子击球手表现优于左撇子击球手时，他们的表现就优于击球手得到的平均适应值：

$$44-14p^t>30+6p^t$$

调整关于右撇子击球手比例的表达式，我们得到 $p^t<0.7$。如果低于 70% 的投手是右撇子，那么右撇子击球手表现优于左撇子击球手，且根据复制者动态，右撇子击球手的比例会增大。总之，这个复制者动态告诉我们的是击球手群体以下列方式演化：

$$\begin{aligned}&如果\ p^t<0.7，那么\ b^{t+1}>b^t；\\&如果\ p^t=0.7，那么\ b^{t+1}=b^t；\\&如果\ p^t>0.7，那么\ b^{t+1}<b^t。\end{aligned} \tag{17.4}$$

接下来，我们分析投手群体的情况。右撇子投手的平均适应值等于：

$$F_P^t（右撇子）=b^t\times 30+(1-b^t)\times 21=21+9b^t$$

左撇子的适应值为：

$$F_P^t（左撇子）=b^t\times 24+(1-b^t)\times 30=30-6b^t$$

由复制者动态得出：

$$p^{t+1}=p^t\times\left[\frac{21+9b^t}{p^t\times(21+9b^t)+(1-p^t)\times(30-6b^t)}\right] \tag{17.5}$$

从这个方程式中，我们得到以下关于投手左右手使

用习惯的描述：

$$\begin{aligned}&\text{如果 } b^t>0.6\text{，那么 } p^{t+1}>p^t\text{；}\\&\text{如果 } b^t=0.6\text{，那么 } p^{t+1}=p^t\text{；}\\&\text{如果 } b^t<0.6\text{，那么 } p^{t+1}<p^t\text{。}\end{aligned}\tag{17.6}$$

如果超过 60% 的击球手是右撇子的话，那么右撇子击球手的比例会逐渐增加。

右撇子击球手的适应值和左撇子击球手的适应值相同的复制者动态出现的静止点需要满足条件 $p^t=0.7$；右撇子投手的适应值和左撇子投手的适应值相同的复制者动态出现的静止点需要满足条件 $b^t=0.6$。当然，这仅仅说明如果群体结构开始时 60% 为右撇子击球手和 70% 为右撇子投手，那么群体结构就是稳定的。但正如我们从图 17—8 中知道的，它一开始的结构并不是这样的；它一开始有 85%～90% 的击球手和投手是右撇子。从这个点开始的话，复制者动态会把群体结构推向静止点吗？

为了探索这个问题，我们将采用图 17—10 中的相图。横坐标表示 b^t 的值，纵坐标表示 p^t 的值。箭头表示了表达式（17.4）和表达式（17.6）中所确定的群体结构的方向。比如，思考Ⅱ象限，它包含所有的 $b^t<0.6$ 和 $p^t>0.7$ 的群体结构。从表达式（17.6）中我们知道，当 $b^t<0.6$ 时，$p^{t+1}<p^t$；就是说，如果有足够少的右撇子击球手，那么左撇子投手的表现优于右撇子投手的表现，且根据复制者动态，右撇子投手的比例在下降。如果 $p^t>0.7$，那么由表达式（17.4）我们知道，右撇子击球手的比例在下降。所以，当群体结构处于Ⅱ象限，如箭头方向所反映的，b^t 和 p^t 的值都在下降。同理，根据表达式（17.4）和表达式（17.6），我们可以确定箭头在其他象限的指向。

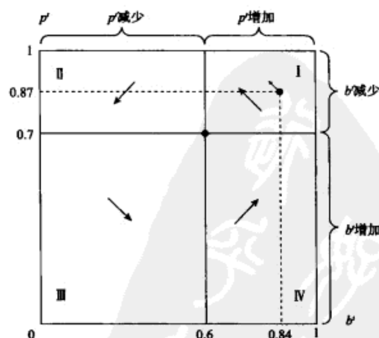


图 17—10　击球手和投手左右手习惯的相图

这个相图是有用的，因为它帮助我们理解棒球员的群体是如何演化的。静止点是（$b^t=0.6$，$p^t=0.7$），重要的是复制者动态会不会把群体状态推到这一点。遗憾的是，情况似乎不是这样的。假设群体刚开始时击球手的 84% 为右撇子，投手的 87% 为右撇子，其实这也是 19 世纪 70 年代美国职棒大联盟的群体比例。那么，这个群体的比例属于第一象限，根据图 17—10，我们知道右撇子击球手的比例会下降，这有利于群体发展，因为它的走势使群体结构逐渐接近静止点的值 60%。但是右撇子投手的比例会增加，这样的话它就远离静止点的值 70%。如果它一直朝着这个轨迹发展，不久情况就成为Ⅱ象限的情况，这样右撇子击球手的比例会下降至 60% 以下，尽管现在投手的比例也在下降。沿着这个轨迹当群体结构的情况成为Ⅲ象限，然后进入Ⅳ象限，我们发现它很难到达静止点。

群体结构是否能演化到静止点？我们可以通过把表达式（17.3）和表达式（17.5）中的复制者动态方程编码到计算机来解答。如果我们把表达式（17.3）和表达式（17.5）中的 b^t 和 p^t 分别替换为 0.84 和 0.87，得出 $b^{t+1}=0.826$ 和 $p^{t+1}=0.884$。所以，在 t 代的下一代中，击球手中 82.6% 为右撇子，投手的 88.4% 是右撇子，正如图 17—10 中箭头所指的。如果我们把 0.826 和 0.884 分别代入表达式（17.3）和表达式（17.5）中的 b^t 和 p^t，我们得到 $b^{t+1}=0.810$ 和 $p^{t+1}=0.896$。以这种迭代方式类推，我们可以确定击球手和投手的左右手习惯是如何演化的，得到的结果如图 17—11 所示。群体结构不断循环且越来越远离静止点。即使群体结构无限接近静止点，又开始新一轮的循环了。对这场博弈来说，单独的静止点不是一个吸引子。

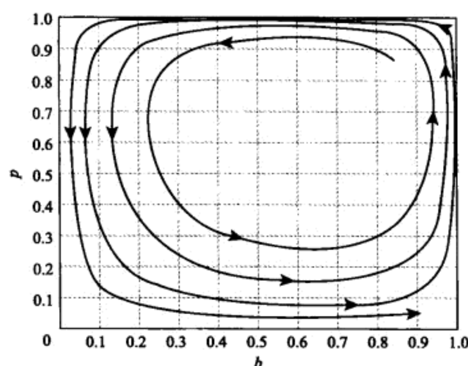


图 17—11　棒球手左右手习惯演化的轨迹

遗憾的是，我们的演化博弈理论模型不可能模拟出与图 17—8 中描述的历史趋势相一致的结果。好吧，我们不可能做什么都成功。但是这并不代表彻底的失败，因为通过学习模型不能解释棒球的左右手用手习惯问题，我们就缩小了模型能解释的范围研究。正如你正在寻找一个丢失的东西，在壁橱里找不到你也许会很沮丧，但至少你知道你必须从其他地方去寻找。这里，我们已经排除了一个模型，现在我们可以把精力放在探索其他模型上。

合作的演化

> 在人类（也包括动物界）发展的历史长河中，那些懂得有效合作与随机应变的往往占据着优势。
>
> ——查尔斯·达尔文

尽管由激烈竞争使得自然选择机制的存在，实际上动物王国里还是存在着大量合作的例子。如果你读了 14 章，你会回忆起来吸血蝙蝠是通过共享食物进行合作的，一只蝙蝠的胃里存有母牛的血来供饥饿的蝙蝠食用。互相梳理毛是动物合作中最常见的形式。比如，黑斑羚用舌头驱赶颈部的虱子。另一个例子是具有感召力的警告。一只动物看见天敌，冒着把其注意力引导自己身上的危险，发出警告来召唤它的同伴。这种示意性的行为也在长尾黑颚猴中存在，长尾黑颚猴已经进化地可以针对三种猎物（鹰、豹和蛇）发出不同的信号，这是很有意义的，因为回应的方式决定了猎物是在空中、地面还是在树上。进化生物学的一个重大挑战就是在自然选择机制的框架下解释这种合作的行为。这方面的研究已取得丰厚的成果，因为生物学家已经发现自然选择机制下合作的方法。让我们在这里探讨其中的一种方法。

分析合作的可能性的最常见的模型是"囚徒困境"，如图 17—12 所示。群体中的两个成员搭档，每个都决定是背叛还是合作。在动物理毛行为中，背叛就是不为其他动物梳理毛，合作就是为其他动物梳理毛。尽管它们意识到都选择合作比都选择背叛获得更高的适应值，问题是背叛严格优于合作。如果这真是它们所面临的环境，那么很明显合作不可能出现。

在许多场景下，合作的例子随处可见，一个关键的特征是动物重复的博弈。例如，黑斑羚有很多机会可以互相理毛。这样，当群体的两个成员博弈，我们假设它们在"囚徒博弈"中博弈 10 次（而不是 1 次）。现在一种策略就是和在 10 次博弈中如何行为相关的一个决策规定。它可以是选择一系列行动的规定，如前 7 次合作，最后 3 次背叛。或者它也可以是一个规定，根据规定采取什么行动视对方的策略而定。比如，一开始合作，但如果另一个参与者从此选择背叛，那么它选择永远背叛（这就成为触发策略，如果你回忆第 13 章的内容）。

		参与者 2	
		背叛	合作
参与者1	背叛	3, 3	8, 1
	合作	1, 8	5, 5

图 17—12　"囚徒困境"

在这 10 次重复"囚徒博弈"中，有许多可行的策略。为了使分析变得简单，我们假设在"基因"库中只有 3 种策略可供选择也就是说，每个成员被赋予使用三种策略。一种策略是永远合作，我们称之为"合作者"，第二种策略是永远背叛，我们称之为"背叛者"，这两种都比较简单，因为一个参与者的行为和其他成员采取的策略无关。第三种策略，我们称之为"以牙还牙"策略，这使得采取的行动取决于之前发生的情况。一个被赋予"以牙还牙"策略的参与者在第一轮选择合作，在其他任何博弈中，都根据对方在前一轮博弈的情况选择自己的策略。因为"以牙还牙"策略是回应性的，那么它奖励选择合作的参与者，惩罚选择背叛的参与者。

是我们想当然地认为动物就是采取"以牙还牙"策略的吗？实际上，在许多物种都会选择"以牙还牙"策略。刺鱼是很小的鱼，当有猎物如梭鱼接近时，它可能向猎物靠近来施加威胁。因为如果有两只刺鱼合作会更安全，合作行为就涉及两只刺鱼去接近猎物，而在前进中，背叛行为使得一只刺鱼落在另一只后面。我们以在一个水箱中放一只刺鱼的实验来检验"以牙还牙"这一策略。再用一面镜子来模仿"另一只刺鱼"的行为（实际上是镜子中这只刺鱼表现的映射）。镜子使得这只刺鱼很清楚"另一只刺鱼"的合作行为（随着这只刺鱼前进），而另一面镜子使得这只刺鱼很清楚"另一只刺鱼"的背叛行为（落在真实刺鱼的后面）。你瞧，刺鱼的行动和"以牙还牙"策略一致：在下一轮博弈中，当镜子映射的前一轮中，"另一只"是合作的，它在这一轮就

选择合作；当镜子映射的前一轮中，“另一只”是背叛的，它在这一轮中就选择背叛。

被赋予“合作”“背叛”和“以牙还牙”策略中的一种策略，群体两个成员进行10次博弈。图17—13描述了博弈的战略形式。如果双方都被赋予背叛策略，那么在10次博弈中，每个成员都选择背叛策略。在每一轮得到的适应值为3，所有回合下来得到的适应值就为30。这就是当两个选择背叛的成员博弈时，我们得到表格中的（30，30）适应值。现在，假设一个选择背叛的和另一个选择合作的博弈。那么在每一轮，前者选择背叛，后者选择合作，所以选择背叛的得到的适应值为80，而选择合作的得到的适应值只有10。最有意思的是至少有一个成员被赋予“以牙还牙”策略的情况。假设采取“以牙还牙”策略的遇见了一个采取合作的参与者。两者在第一轮中都选择合作。在第二轮中，先前采取合作的仍然选择合作策略（因为它知道如何做）之前被赋予“以牙还牙”策略的参与者选择合作策略，因为它要做出相同的回应。就这样在所有的博弈中一直持续，所以每个参与者得到的适应值为50。当一个被赋予“以牙还牙”策略的参与者和另一个被赋予“以牙还牙”策略的参与者博弈，出现的结果和上一个情况相同。现在，如果一个被赋予“以牙还牙”策略的参与者和另一个被赋予“背叛”策略的参与者博弈，结果就会完全不同。在第一轮，被赋予“以牙还牙”策略的参与者选择合作策略，而被赋予背叛策略的选择背叛策略。作为报复，被赋予“以牙还牙”策略的参与者在第二轮中就会选择背叛策略，被赋予背叛策略的依然选择背叛策略。因为被赋予背叛策略的总是选择背叛策略，被赋予“以牙还牙”策略的参与者作为回应也选择背叛策略。所以，被赋予“以牙还牙”策略的参与者得到的适应值为28（在第一轮中得到1个适应值，其他的九次每次都为3），被赋予背叛策略的参与者得到的适应值为35（第一次为8，后面9次都为3）。

参与者2

参与者1		背叛者	合作者	以牙还牙
	背叛	30, 30	80, 10	35, 28
	合作	10, 80	50, 50	50, 50
	以牙还牙	28, 35	50, 50	50, 50

图17—13 “重复囚徒困境”

通过对这些策略进行比较，我们知道没有任何一种策略是严格劣势策略。如果另一个参与者选择“以牙还牙”策略，那么这个参与者采取合作策略还是采取“以牙还牙”策略都是最好的结果。而对背叛者的最好回应就是也做一个背叛者。然而，注意“以牙还牙”策略弱占优于合作策略：不管是遇到采取合作策略的还是采取“以牙还牙”策略的，他们得到的适应值都一样。但是当遇见的是背叛者，采取“以牙还牙”策略就优于采取合作策略，因为采取“以牙还牙”策略的参与者只在第一轮中“吃亏”，而采取合作策略的10次都会被背叛者敲竹杠。

我们希望描述群体的混合策略随着时间是如何演化的。一个群体最终会被背叛者主导，所以每个成员是否都对彼此讨厌？或者合作者和“以牙还牙”者大量存在使得社会是合作性的？

利用复制者动态解决这些问题，我们就需要推导采取这三种策略得到的适应值。在t代中，令p_C^t表示群体中合作者的比例，p_T^t表示群体中“以牙还牙”者的比例，p_D^t（$=1-p_C^t-p_T^t$）表示群体中背叛者的比例。那么采取每种策略得到的平均适应值为以下：

采取合作策略得到的适应值：
$$p_C^t\times 50+p_T^t\times 50+p_D^t\times 10$$
采取“以牙还牙”策略得到的适应值：
$$p_C^t\times 50+p_T^t\times 50+p_D^t\times 28 \tag{17.7}$$
采取背叛策略得到的适应值：
$$p_C^t\times 80+p_T^t\times 35+p_D^t\times 30$$

思考采取“以牙还牙”策略的例子。一个“以牙还牙”者遇见一个合作者的概率为p_C^t，得到的适应值为50；遇见另一个“以牙还牙”者的概率为p_T^t，得到的适应值为50；遇见一个背叛者的概率为p_D^t，得到的适应值为28。把表达式（17.7）中的p_D^t替换为$1-p_C^t-p_T^t$，再简化表达式，得到：

采取合作策略得到的适应值：
$$10+40p_C^t+40p_T^t$$
采取“以牙还牙”策略得到的适应值：
$$28+22p_C^t+22p_T^t \tag{17.8}$$
采取背叛擦亮得到的适应值：
$$30+50p_C^t+5p_T^t$$

用这三种策略分析复制者动态具有很大的挑战性。所以，我们还是把复制者动态方程写下来，尽管这很复杂。我们将用相图来分析，构建相图的第一步是估算每种策略在和其他策略博弈时的相对适应值。首先“以牙还牙”者和合作者之间的博弈结合表达式（17.8）分析，我们发现当且仅当：

$28 + 22p_C^t + 22p_T^t > 10 + 40p_C^t + 40p_T^t$，或者，$p_C^t + p_T^t < 1$

时，前者策略更优。如果至少存在一些背叛者（使得 $p_C^t + p_T^t < 1$），那么采取“以牙还牙”策略得到的适应值就比采取合作策略得到的适应值高。然而，如果 $p_C^t + p_T^t = 1$，那么采取“以牙还牙”策略和采取合作策略得到的适应值一样多。记住，只有当和背叛者博弈时，采取这两种策略时最终得到的适应值相同。

那么，当 $p_C^t + p_T^t < 1$ 时，采取“以牙还牙”策略比采取合作策略得到的适应值更高，且当 $p_C^t + p_T^t = 1$ 时，采取“以牙还牙”策略和采取合作策略得到相同的适应值。

接下来，我们比较采取“以牙还牙”策略得到的适应值和采取背叛策略得到的适应值。当：

$28 + 22p_C^t + 22p_T^t > 30 + 50p_C^t + 5p_T^t$，

或者当 $p_T^t > \dfrac{2+28p_C^t}{17}$ 时，

采取“以牙还牙”策略更优。如果 $p_T^t > \dfrac{2+28p_C^t}{17}$，那么采取“以牙还牙”策略胜于采取背叛策略。当遇见另一个参与者也采取“以牙还牙”策略时，这个采取“以牙还牙”策略的参与者就会有一系列合作的机会，而另一个采取背叛策略的参与者只在第一轮中得到高的适应值。要使“以牙还牙”策略优于背叛策略，还需要满足的条件是不要有太多的参与者采取合作策略，因为尽管采取“以牙还牙”策略的参与者和采取合作策略的参与者合作，相对那些天真的采取合作策略的参与者来说，采取背叛策略的参与者得到的适应值更高。当：

$$p_T^t > (<) \frac{2+28p_C^t}{17} \tag{17.9}$$

时，采取“以牙还牙”策略得到的适应值高于（低于）采取背叛策略得到的适应值。

让我们从这条信息开始构建图 17—14 的相图。相图中的每一点表示一个特殊的群体结构。例如，点 $(p_T^t, p_C^t) = (0.3, 0.5)$ 代表的群体结构为：在群体中，30% 采取的是“以牙还牙”策略，50% 采取的是合作策略，其他 20% 采取的是背叛策略。注意 $(p_T^t, p_C^t) = (0,0)$ 表示群体结构中所有参与者采取背叛策略，且在（1，0）和（0，1）直线上的所有点满足 $p_T^t + p_C^t = 1$，所以在这类群体中没有背叛者。

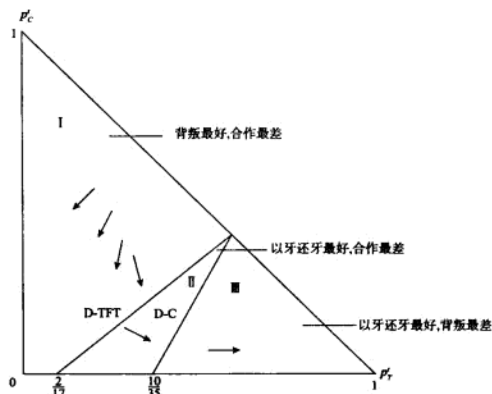


图 17—14　“重复囚徒博弈”的相图

由（17.9）可知，$\dfrac{2+28p_C^t}{17}$ 是确定采取背叛策略或者“以牙还牙”策略能否获得更高适应值时 p_T^t 的临界值。当 $p_C^t=0$ 时，$\dfrac{2+28p_C^t}{17} = \dfrac{2}{17}$，且当 p_C^t 增大时，$\dfrac{2+28p_C^t}{17}$ 也增大，所以 D–TFT 就是一条向上的直线。位于 D–TFT 线左边区域的群体结构是这样的结构，其中背叛者得到的适应值高于“以牙还牙”者。因为我们也已经证明了“以牙还牙”者得到的适应值高于合作者得到的适应值，那么当群体结构位于 D–TFT 线左边区域时，背叛者得到的适应值最高。如果群体结构处于这个区域（表示为 I 象限），那么背叛者的比例会越来越大（因为背叛者得到的适应值最高，所以一定大于群体的平均适应值），同时，合作者的比例越来越小（因为合作者得到的适应值最低，所以一定低于群体的平均适应值）。因为背叛者逐渐增多，群体结构会像“南”移动；它也可能向“西南”向移动（这样的话，“以牙还牙”者和合作者的比例都在下降）；或者向“东南”方向移动（这样，合作者的比例逐步缩小，但“以牙还牙”者的比例在增大）。I 象限中箭头就表明了群体结构的走势。

当群体结构处于 D–TFT 线的右侧时，“以牙还牙”

者得到的适应值高于背叛者得到的适应值。所以，“以牙还牙”策略是最优策略，因为它总是优于合作策略。于是，当群体结构处于D–TFT线的右侧时，被赋予“以牙还牙”策略的参与者会越来越多。为了确定背叛者和合作者孰优孰劣，我们需要比较他们的相对适应值。当

$$10 + 40p_C^t + 40p_T^t > 30 + 50p_C^t + 5p_T^t$$

时，合作者得到的适应值高于背叛者得到的适应值。解关于p_T^t的不等式，我们得到：

$$\text{当 } p_T^t > (<) \ \frac{20+10p_C^t}{35} \qquad (17.10)$$

时，合作者得到的适应值高于（低于）背叛者得到的适应值。由不等式（17.10）得出，如果群体中有足够多的“以牙还牙”者，那么合作者得到的适应值高于背叛者得到的适应值。在和“以牙还牙”者博弈时，合作者得到的适应值更高（由于他们合作广泛）而背叛者得不到这么多，因为他只在第一轮和“以牙还牙”者博弈时得到高的适应值。然而，当群体中有许多合作者时，背叛者比合作者更适合在群体中存在。尽管当和另一个采取合作策略的参与者博弈时，合作者做得更好，但背叛者在一个群体中表现得更好，如果这个群体中包括许多只知道自己如何才能做得很好的参与者。

将从不等式（17.10）得出的关键临界值 $\frac{20+10p_C^t}{35}$ 在图17—14中标出，且是D–C线的标志点。当群体结构处于D–C线的左侧时，背叛者得到的适应值高于合作者得到的适应值；当群体结构处于D–C线的右侧时，合作者得到的适应值高于背叛者得到的适应值。所以，对处于Ⅱ象限的群体结构来说，“以牙还牙”者得到的适应值高于背叛者得到的适应值，而背叛者得到的适应值高于合作者得到的适应值。于是，群体中的“以牙还牙”者越来越多，合作者这个子群体的规模越来越小；群体结构朝着“东南”方向移动（注意箭头的方向）。最后，如果群体结构处于Ⅲ象限，位于D–TFT线和D–C线的右侧，这样的话，“以牙还牙”者得到的适应值高于合作者得到的适应值，而合作者得到的适应值高于背叛者得到的适应值。所以，在Ⅲ象限，“以牙还牙”者的比例越来越大，背叛者的比例越来越小。

接下来就是利用相图来预测群体演化的走势和吸引子。假设群体刚开始有很少成员被赋予“以牙还牙”策略，如图17—15所示。那么背叛者得到的适应值最高，原因如下：如果群体中有许多合作者，那么即使“以牙还牙”者能和他们保持合作，背叛者在和合作者博弈时得到的适应值比“以牙还牙”者和合作者博弈得到的适应值更高。如果群体中有很少的合作者，那么就没有足够的合作者与“以牙还牙”者进行博弈，同样，背叛者得到的适应值更高。群体中背叛者的数量稳定增加，合作者和“以牙还牙”者被淘汰。一个吸引子就是群体中的全体成员采取背叛策略，这样在群体中就没有合作出现。只有当群体刚开始“以牙还牙”者的数量很少时才会发生这样的情况。

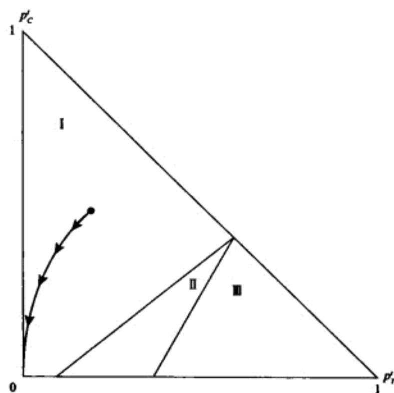


图17—15　群体中“比牙还牙”者比例很低时群的走势

现在考虑图17—16中群体中刚开始的比例情况，在群体中，有足够多的成员采取“以牙还牙”策略，很少的成员采取合作策略。尽管背叛者得到的适应值仍然最高（这样它的比例就一直增加），但同样，“以牙还牙”者的比例也在增加。“以牙还牙”者有许多可以合作的成员，所以它会理性地选择采取最好的行动。然而，合作者的比例一直降低，因为他们被背叛者欺负。因为群体中“以牙还牙”者的比例一直增大，合作者的比例一直减少，这个环境越来越有利于“以牙还牙”者，对背叛者越来越不适宜。当群体结构进入Ⅱ象限，“以牙还牙”者得到的适应值最高，所以它的比例持续增加，而合作者比例持续减少。这种情况导致群体结构进入Ⅲ象限，现在群体中背叛者得到的适应值最低，因为群体中可以被欺负的合作者的数量太少了。背叛者这个子群体的数量就逐渐减少，而“以牙还牙”者和合作者的比例

就逐渐增大。尽管最终群体的大部分是“以牙还牙”者，但仍然有一些有合作者特征的子群体存在。最后，群体中的大部分参与者是合作者。

对那些担心人类未来的人来说，了解了合作是演化稳定的可能是一种慰藉。如果群体结构一开始由具有“以牙还牙”特征的成员组成，那么这种特征会扩散很广，因为“以牙还牙”者可以和同类（还有合作者）保持稳定的合作。尽管背叛者和背叛者博弈的结果优于和“以牙还牙”者博弈的结果，但得到的适应值还是低于“以牙还牙”者和“以牙还牙”者博弈得到的适应值。

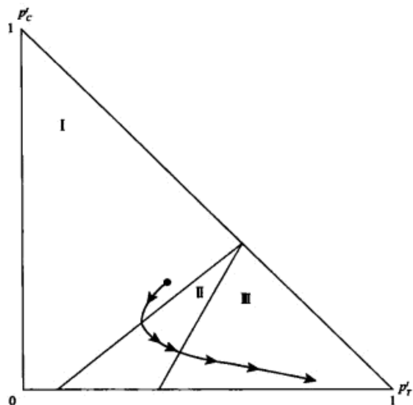


图 17—16　群体中以牙还牙者比例很高时群体的走势

注意，如果群体中有许多的合作者（这种情况属于Ⅰ象限），那么合作者被淘汰，因为背叛策略是占优策略。当群体中有很多“以牙还牙”者，但不是很多的合作者，合作才会出现。因为“以牙还牙”者很快对背叛者采取报复行为，而合作者就任凭背叛者一次又一次利用自己。要使合作行为一直延续，紧靠合作者群体是不够的；参与者也一定得随时惩罚背叛者，因为只有这样背叛者才会被群体淘汰。

尽管合作可以发生，一个有趣的因素可以扰乱群体合作的稳定性。假设所有的成员都采取“以牙还牙”策略，且有一部分突变体采取合作策略。合作策略不可能被淘汰，因为合作者得到的适应值和“以牙还牙”者得到的适应值相同。所以，若群体大部分是“以牙还牙”者，而只有一些合作者，群体稳定状态可以持续。如果突变体的数量增加，群体结构将“漂浮”在 $p'_T+p'_C=1$ 直线附近，如图 17—17 所示。然而，我必须强调一点就是没有因素可以使群体结构朝那个方向演化。而如果有一些“以牙还牙”者突变为合作者，那么群体结构将朝那个方向演化。

当群体中有很多合作者，群体结构会发生什么变化呢？群体结构将进入Ⅰ象限，但只是在边缘移动。现在假设群体引入一些背叛者，那么群体的结构将进入Ⅰ象限内部。这样，背叛者大量增加，因为他们击退所有的合作者。最后，背叛者主导群体结构。这样的话，合作就不稳定。

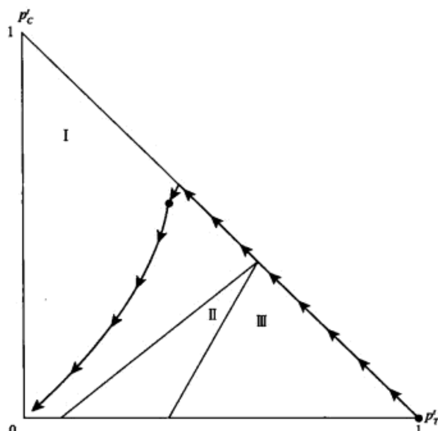


图 17—17　合作的不稳定性

有一个有趣的故事可以解释刚刚描述的动态演化路径。假设没有自然选择机制，而是模仿机制：一种可以得到更高适应值的策略被越来越多的参与者使用，因为其他成员看到采取这个策略可以得到更高的适应值，他们就模仿采取这个策略。现在，如果所有成员都采取“以牙还牙”策略，那么观察到的博弈结果就是他们互相合作。只有当和背叛者博弈时，才可以区分“以牙还牙”者和合作者。但因为群体中没有背叛者，所以就无法区分“以牙还牙”者和合作者的行为。于是，当群体中有新成员加入时，他们会错把“以牙还牙”者当合作者看待，因为他们所看到的总是合作行为。他们意识不到如果另一个参与者采取背叛策略，一个人就应该停止合作。群体就会忘记怎样使群体保持合作（就是说，怎样才能淘汰背叛者）。如果很多人选择合作策略，那么在那个点上，背叛者就会胜过合作者，再进入群体的新成员就会看到，采取背叛策略比采取合作策略得到更高的适应值，所以他们也采取背叛策略而不是合作策略。这样的话，合作规范就被打破，社会由利他主义陷入利己主义。

本章总结与练习题

本章总结

在这章里，我们研究了动物群体（可能是鬣狗、长尾黑颚猴或者是棒球运动员）以及作为模拟的战略博弈的战略场景。群体中的成员随机搭配完成博弈。因为我们不能假定动物具有选择策略的能力，所以假设每个成员被赋予一种策略。尽管有一些动物不能选择如何博弈，但群体通过自然选择机制“选择”那些更适合生存的策略。那些更优的策略——量化为适应值的高低，在群体的下一代占有更大的份额。

演化博弈理论发展的一个自然选择机制模型是**复制者动态**，它规定了当采取一种策略得到的适应值高于（低于）群体的平均适应值时，群体中采取这种策略的成员就增加（减少）。借助复制者动态，本章探讨了当成员采取混合策略时，复制者动态把群体演化到那个阶段。

自然选择机制使得复制者动态到达的最后点称为**静止点**。一个静止点是一个群体结构比例，一旦群体结构比例发展到那个点，复制者动态就使群体稳定在这个点。一个静止点是局部稳定的，如果当群体结构接近那个静止点，复制者动态会推动群体到那个静止点。一个局部稳定的静止点被称为**吸引子**。吸引子可以有很多，群体一开始的结构比例决定了吸引子。那些所有群体结构的开始比例可以引导群体到达吸引子的区域被称为**吸引域**。为了探索复制者动态下的群体如何演化，我们引入了相图的概念。由于一个相图需要许多语言来解释，我建议你通过复习图17—10来了解。（好吧，我承认我比较懒，但你应试着用语言来描述相图。）

借助这些概念，我们探索了群体演化为共同捕猎或者单个捕猎时的情况。我们借助猎鹿博弈的模型来研究，猎鹿博弈在演化博弈理论之前的将近两个世纪已经发展成熟。从更宽泛的角度来看合作问题，我们分析了“重复囚徒困境”。通过“重复囚徒困境”，我们发现群体演化的合作规范，那就是，合作发生在吸引子点上，从一开始群体中有足够多的成员被赋予“以牙还牙”策略的时候开始。“以牙还牙”策略的关键特征是和对方做出相同的举动，所以如果另一方合作，他就合作，而如果另一方背叛，那么他也背叛。相反，如果群体中的成员总是采取合作策略，那么具有讽刺意义的是，合作者被淘汰。问题是在这么多合作者存在的群体中，背叛者占主导地位。要使社会保持合作规范，每个人不仅必须做好事，而且要能惩罚那些背叛者。一个群体必须有适度的斗争和合作，这是博弈理论最重要的一课。

练习题

1. 思考图17—18所示的博弈，令 x^t 表示在 t 代中群体被赋予 x 策略的比例。

a. 根据复制者动态，描述群体如何演化。

b. 找到吸引子和每个吸引子的吸引域。

参与者 2

参与者 1		x	y
	x	4,4	9,2
	y	2,9	10,10

图 17—18

2. 思考图17—19中的博弈，令 x^t 表示在 t 代中群体被赋予 x 策略的比例。

a. 根据复制者动态，描述群体如何演化。

b. 找到吸引子和每个吸引子的吸引域。

参与者 2

参与者 1		x	y
	x	3, 3	2,6
	y	6, 2	1, 1

图 17—19

3. 思考图 17—20 中的博弈，令 a^t 和 b^t 分别表示 t 代中群体被赋予策略 a 和策略 b 的比例。所以 $1-a^t-b^t$ 就是被赋予策略 c 的比例。

a. 列出复制者动态方程。

b. 画出相图，横坐标表示 a^t 的值，纵坐标表示 b^t 的值。

c. 找出所有的静止点。

参与者 2

		a	b	c
参与者 1	a	4, 4	2, 4	1, 2
	b	4, 2	2, 2	1, 3
	c	2, 1	3, 1	5, 5

图 17—20

4. 思考图 17—21 中的博弈。

a. 列出复制者动态方程。

b. 找到吸引子和每个吸引子的吸引域。

参与者 2

		低	中	高
参与者 1	低	2, 2	2, 3	4, 1
	中	3, 2	4, 4	5, 3
	高	1, 4	3, 5	4, 4

图 17—21

5. 回到 17.4 小节中 10 次“重复囚徒困境”。现在假设仅仅有两种策略：背叛策略和“以牙还牙”策略。

a. 根据复制者动态，描述群体如何演化。

b. 找到吸引子和每个吸引子的吸引域。

6. 回到 17.4 小节中 10 次“重复囚徒困境”。现在假设有三种策略：（1）背叛策略；（2）“以牙还牙”策略；（3）刁钻以牙还牙策略。前面两种策略已经介绍过了，而第三种策略和“以牙还牙”策略很相似，但不同的是无论对方在第 9 次采取什么策略，它在第 10 次总是选择背叛策略。让 p_S^t 和 p_T^t 分别代表在 t 代中被赋予刁钻以牙还牙策略和“以牙还牙”策略的比例。

a. 推导适应值矩阵（模拟图 17—13）。

b. 比较每种策略的适应值。

c. 群体中所有成员被赋予“以牙还牙”策略时的结构是不是一个吸引子？

d. 群体中所有成员被赋予背叛策略时的结构是不是一个吸引子？

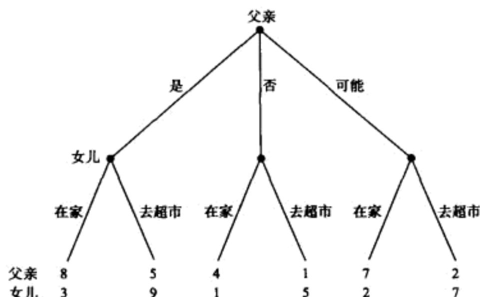
e. 群体中所有成员被赋予刁钻以牙还牙策略时的结构是不是一个吸引子？

7. 返回到图 17—5 中的猎鹿博弈。我们已经证明群体中有$\frac{2}{3}$的成员被赋予鹰策略时是一个静止点，但不是一个吸引子。现在思考一个同质群体中所有成员采取混合策略，在混合策略中选择鹰策略的概率为$\frac{2}{3}$。证明这个策略不是一个演化稳定策略。

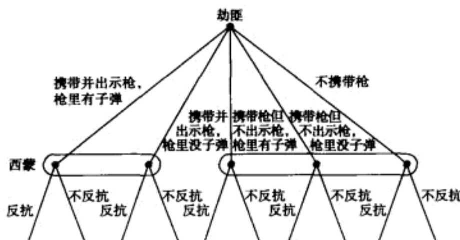
理解测验答案

第 2 章

2.1



2.2



2.3 因为薇薇卡仍然只有一个信息集，那么薇薇卡的策略集就与推导出的图 2—1 扩展式博弈的策略集相同，她仍有两种策略：（1）如果盖伊实施了绑架，那么薇薇卡付赎金；（2）如果盖伊实施了绑架，那么薇薇卡不付赎金。相反，盖伊的策略模版有所改变，因为现在他只有 2 个信息集：

在首节点，______[填入绑架或不绑架]。

如果盖伊实施了绑架，那么______[填入撕票或释放]。

在这两个信息集中，每个信息集有 2 种可行行动方案，所以盖伊有四种策略：（1）在首节点，绑架，如果实施了绑架，那么撕票；（2）在首节点，绑架，如果实施了绑架，那么释放奥兰多；（3）在首节点，不绑架，如果实施了绑架，那么撕票；（4）在首节点，不绑架，如果实施了绑架，那么释放奥兰多。尽管没有改变盖伊的可供选择的行动方案，但信息集的改变影响了他的策略集。

2.4 劫匪有 3 种策略：持枪，且出示手枪；持枪，不出示手枪；不携带手枪。西蒙有 2 个信息集，所以对他来说，一个策略包括两种行动方案：如果劫匪出示手枪，他该如何采取行动，以及如果劫匪不出示手枪如何采取行动。于是，西蒙有四种策略：*R/R, R/DNR, DNR/R,* 和 *DNR/DNR*（这里 *R* 代表反抗，*DNR* 代表不反抗，第一个行动方案是指当劫匪出示手枪时的信息集）。策略式如下：

劫匪 \ 西蒙	*R/R*	*R/DNR*	*DNR/R*	*DNR/DNR*
携带且出示手枪	3,2	3,2	4,5	4,5
携带但不出示手枪	3,2	5,4	3,2	5,4
没携带手枪	2,6	6,3	2,6	6,3

第 3 章

3.1 对参与者 1 来说，策略 *a* 对 *c* 严格占优，*b* 对 *d* 弱占优；对参与者 2 来说，策略 *y* 对 *w* 和 *x* 严格占优，*w* 对 *x* 弱占优。

3.2 对参与者 1, 来说，策略 *a* 对 *c* 严格占优；对参与者 2 来说，策略 *y* 对 *w* 和 *x* 严格占优。因为参与者是理性的，他们的策略可被排除。由于参与者 1 知道参与者 2 是理性的，因而参与者 1 随即知道参与者 2 不会使用 *w* 和 *x* 策略。在排除了参与者 2 这些策略后，从参与者 1 角度来说，策略 *a* 对 *b* 严格占优。以此类推，由于参与者 2 知道参与者 1 是理性的，参与者 2 即知参与者 1 不会选择策略 *c*。在

排除了参与者 1 这些策略后，从参与者 2 角度来说，策略 z 对 y 严格占优。则答案为参与者 1 选择策略 a 和 d，参与者 2 选择策略 z。

3.3 由于此博弈是在理解测验 3.2 中使用的同一个博弈，我们可以接着讨论步骤 3。在步骤 2 之后，参与者 1 还有策略 a 和 d，参与者 2 还有策略 z。策略 d 对 a 严格占优。因此，此博弈是占优可解的，在重复剔除严格劣势策略中存在唯一的策略：参与者 1 选 d 且参与者 2 选 z。

第 4 章

4.1 存在三个纳什均衡解：（b，x），（a，z）和（d，z）。

4.2 存在两个纳什均衡解：（c，y，Ⅰ）和（a，x，Ⅱ）。

第 5 章

5.1 令（x，y，z）代表有 x 个星肚史尼奇、y 个横条史尼奇和 z 个光肚史尼奇的策略组合。总共 7 个史尼奇，我们从（0,1,6）开始分析。这不是一个纳什均衡，因为光肚史尼奇收益为 0（横条史尼奇受到较少的尊敬），但可以通过纹星或者纹横条获得收益 1。再看（0,2,5），一个光肚史尼奇可以通过纹星或横条获得收益的提升；前两种情况下，星肚和横条都是在最小的群体中。（0,3,4），一个光肚史尼奇通过纹星将自己的收益从 0 提升为 1。（注意，尽管如此，纹横条是没什么用的）（1,1,5）是一个纳什均衡，因为星肚史尼奇和横条史尼奇各有 1 个收益，且不会变得更好，同时光肚史尼奇将仍非最小群体，就算有一个将要纹星或纹横条去。（1,3,3）不是纳什均衡，因为一个横条史尼奇可以反过来纹一个星而获得更好受益。最后，（2,2,3）是纳什均衡。我们可以总结出达到纳什均衡需要：（1）一个星肚史尼奇和一个横条史尼奇，（2）2 个星肚史尼奇和 2 个横条史尼奇。还有其他仅仅通过调换位置获得的可比的均衡。一般来说，纳什均衡具有（1）一个单策略的史尼奇，一个有第二策略的史尼奇以及其他有第三策略的史尼奇；（2）两个单策略的史尼奇，两个有第二策略的史尼奇以及其他有第三策略的史尼奇。

5.2 一条航线的纳什均衡选择策略 7，其他 n–1 条航线的纳什均衡选策略 1。一开始，考虑在 $\{s_1,\cdots,s_n\}$ 中选择最大值的航线，设为航线 1。然后其收益为 $50+10\times\max\{s_1,\cdots,s_n\}$。若 $\max\{s_1,\cdots,s_n\}<7$，则航线 1 选择策略 7 更好，并获得 120 的收益。因此，在纳什均衡中，必然是 $\max\{s_1,\cdots,s_n\}=7$。现在假定 2 条或更多航线选择策略 7，则其中的任何一个都可能将其策略降为 1，将收益从 120 提高到 180。我们因此得出结论，只能有一条航线选择策略 7，正如刚刚讨论过的，没选择策略 7 的航线将会选择策略 1.

5.3 任何一个三家公司进入市场的策略组合都是纳什均衡，其他任何策略组合均不是。从表 5—6 中可以看到，当其他两家公司进入时，第三家公司的进入是有利可图的。因此，只有两家公司进入（或者，十分显而易见的，只有一家公司进入或没有公司进入）并不能达到均衡。我们同样注意到，当三家公司进入市场时，对任何第四个想进入的公司来说是无利可图的。因此，全部四家公司进入市场不是纳什均衡。综上所述，如果三家进入，则每个进入者可以获得正收益，所以其进入的决策是最优的；剩下没有进入的公司如果还想进入将得到负收益，所以其不进入的决策也是最优的。

5.4 回忆一下存在的三个纳什均衡：（1）没人抗议，（2）激进分子与进步分子抗议以及（3）人人抗议。每种类型的市民得到的联合收益如下表所示：

来自不同纳什均衡的收益

市民类型	没人抗议	激进分子与进步分子抗议	人人抗议
激进分子	0	4 000	19 000
进步分子	0	2 000	17 000
中产阶级	0	0	5 000

从表中可知，每种类型的市民选择第三种均衡的收益都严格高于选择第一种或第二种均衡的收益。因此，唯一的收益占优纳什均衡解为第三种。

第 6 章

6.1 任何一个三家商店定价至少达到 10 时且至少 2 家商店定价为 10 的策略组合都是纳什均衡，而其他策略组合则不是。假设三家商店定价都高于 10，其中一家高于另外两家，则这家商店具有零需求和零利润。通过轻微削减达到最低价的商店情况会好些，因为这种情况下会有正需求，且由于价格高于成本还有正利润产生。从这个讨论中可以总结出，如果所有的商店定价都高于成本，他们一定会选择相同的定价。但是这也并非一个均衡，因为一个商店能

通过稍稍低于其他两个商店的降价行为获得三倍需求(以及几乎是三倍的利润)。然后我们可以总结出，均衡不能令三家商店定价都高于 10。显然，没有哪家商店愿意定价低于 10，所以假设一家商店定价为 10，此时利润为 0（尽管面向所有消费者）。但是可以通过将价格定为略微低于其他两家中的最低者而获得正利润。现在假定有两家商店定价为 10 了。每家都是零利润，但是一旦定价高于 10，就会有零需求（以及零利润），因为另外一家商店定价为 10。此时，定价高于 10 的商店是零需求，除非它把价格定为 10，但那时仍然是零利润。所以，三家商店的均衡产生了，定价均为 10。

6.2 源自戴尔的最佳反应，得到其遵守自身定价的收益函数的最初衍生公式，并令其等于 0：

$$\frac{\partial V^{\text{Dell}}(P^{\text{Dell}}, P^{\text{HP}})}{\partial P^{\text{Dell}}}=100-2P^{\text{Dell}}+P^{\text{HP}}-2P^{\text{Dell}}+40=0$$

$$=140-4P^{\text{Dell}}+P^{\text{HP}}=0$$

解方程，求出 P^{Dell}，可以看出戴尔公司的最佳反应函数为：

$P^{\text{Dell}}=\dfrac{140+P^{\text{HP}}}{4}$ 或 $\text{BR}^{\text{Dell}}=35+0.25P^{\text{HP}}$

由对称性得，惠普公司具有同样的反应函数：

$\text{BR}^{\text{HP}}=35+0.25P^{\text{Dell}}$

此价格组合（$\hat{P}^{\text{Dell}}$，$\hat{P}^{\text{HP}}$）为当两种计算机产品同时选择最佳反应时的纳什均衡解：

$\hat{P}^{\text{Dell}}=35+0.25\hat{P}^{\text{HP}}$

$\hat{P}^{\text{HP}}=35+0.25\hat{P}^{\text{Dell}}$

将第二个方程代入第一个方程，得到：

$\hat{P}^{\text{Dell}}=35+0.25\times(35+0.25\hat{P}^{\text{Dell}})$

求解戴尔的价格，可得：

$$\hat{P}^{\text{Dell}}=35+8.75+0.625\times\hat{P}^{\text{Dell}}$$

$0.937\,5\times\hat{P}^{\text{Dell}}=43.75$

$\hat{P}^{\text{Dell}}=46.67$

将戴尔的价格四舍五入至小数点后两位，将此数值代入惠普的最佳反应函数，得到：

$\hat{P}^{\text{HP}}=35+0.25\times 46.67$

$\hat{P}^{\text{HP}}=46.67$

得到一个唯一的纳什均衡解，此时两个公司的价格均为 46.67。

第 7 章

7.1 令 p 和 q 分别表示参与者 1 选择上和参与者 2 选择左的概率。令参与者 1 选择上和选择下获得的期望收益相等，有：

$q\times 8+(1-q)\times 1=q\times 3+(1-q)\times 4$

求解，得 $q=3/8$。

令参与者 2 选择左和选择右获得的期望收益相等，有：

$p\times 2+(1-p)\times 5=p\times 6+(1-p)\times 1$

求解，得 $p=1/2$。因此有唯一解。在 $p=1/2$ 且 $q=3/8$ 时有纳什均衡。

7.2 运用重复剔除严格劣势策略，对参与者 1 来说 b 对 a 严格占优，对参与者 2 来说 y 对 z 严格占优。在第二循环阶段，对参与者 2 来说 x 对 w 严格占优。参与者 1 有 b 和 c 两个策略，参与者 2 有 x 和 y 两个策略。简化后的博弈为：

		参与者 2	
		x	y
参与者 1	b	4,1	2,5
	c	2,6	5,2

此博弈有唯一的混合策略纳什均衡解，其中参与者 1 选 b 和选 c 的概率各为 1/2，参与者 2 选 x 的概率为 3/5，选 y 的概率为 2/5。

第 8 章

8.1 在参与者 2 的决策节点上，其最优选择为 x，因为这一选择会产生收益 4，反之选 y 就只能获得收益 1。对参与者 3 的决策节点来说，他会选 d。现在讨论一下参与者 1 的决策节点。通过逆向归纳法，参与者 1 预期从选择 a 中获得收益 1，因为参与者 2 将以选 x 作为回应。若参与者 1 选了 b，结果是收益为 2，因为参与者 3 将随之选择 d。因此，参与者 1 的最优选是 b，唯一的子博弈精炼纳什均衡，即参与者 1 选择 b，参与者 2 选择 x，参与者 3 选择 d。

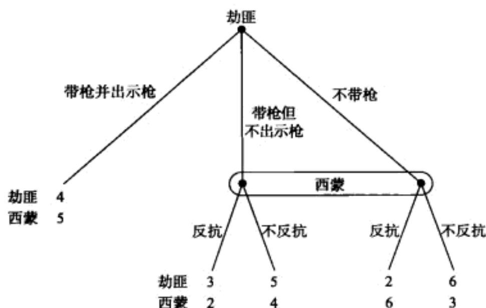
8.2 考虑到参与者 1 的最终决策点，她的最优决策是 y。下一步，考虑基于参与者 1 选择了 R 之后参与者 2 的决策节点，他可以选择 a 获得收益 1 或是选择 b 获得收益 2，因为参与者 1 将会选择 y 来回应。因此，参与者 2 选择 b。在参与者 2 的其他决策节点上，他会选择 b。当参与者 1 在首节点时，她会选 L 并获得收益 3，选 M 获得收益 1（因为参与者 2 会随

之选 b），或者是选 R 获得收益 4（参与者 2 将会选 b 且参与者 1 随后将会选 y）。参与者 1 的最优策略则是 R，其伴随着一个唯一的子博弈精炼纳什均衡就是参与者 1 选择 R 和 y，参与者 2 在两个决策点都选择 b。

第 9 章

9.1 有 7 个子树，但只有 3 个标准子树：a 分支和 x 分支组成的子树；b 分支和 y 分支组成的子树；以及这个树本身。

9.2 思考与劫匪相关的最终的子博弈，在这个子博弈中，劫匪带枪并出示枪。这是一个只有一个参与者的博弈，在这个博弈中存在一个纳什均衡：不反抗。用均衡收益来替换这个子博弈，利用逆向归纳法，我们得到下面的这个博弈：



这个博弈的子博弈就是这个博弈本身。（注意这个树的部分从首节点开始，只包含一个分支，即：带枪但不出示枪。不带枪不是一个子树。这是因为子树不仅包括子博弈开始的节点，还包括子博弈的其他节点，而不带枪这个子树不包括与带枪并出示枪相关的终端节点。）为了求解这个博弈的纳什均衡，我们推导出以下战略式：

西蒙

		反抗	不反抗
劫匪	带枪并出示枪	4,5	4,5
	带枪但不出示枪	3,2	5,4
	不带枪	2,6	6,3

注意西蒙的策略是指他对劫匪带枪但不出示枪行动的反应。这个博弈存在一个纳什均衡（带枪并出示手枪，反抗）

总之，存在一个子博弈精炼纳什均衡，这个均衡是劫匪带枪并出示枪，而西蒙不反抗。

第 10 章

10.1 给定尔普的行动是开枪，很明显，对陌生人来说，（开枪，开枪）是最优的策略。如果他是一个熟练的枪手，他采取开枪策略得到的收益为 3，采取僵持策略得到的收益为 1；如果他是一个普通的牛仔，那么他采取开枪策略得到的收益为 2，采取僵持策略得到的收益为 1。所以，对陌生人的任何类型来说，这都是最优策略。现在，尔普的策略是不是最优策略呢？给定陌生人的策略，尔普从开枪策略得到的期望收益为 $0.75\times2+0.25\times5=2.75$。陌生人是熟练枪手的概率为 0.75，尔普开枪得到的收益为 2；陌生人是普通牛仔的概率为 0.25，尔普开枪得到的收益为 5。记得普通牛仔也会采取开枪策略，因为他的策略是（开枪，开枪）。相反，如果尔普僵持，他得到的期望收益为 $0.75\times1+0.25\times6=2.25$。陌生人是熟练枪手的概率为 0.75，尔普得到的收益只有 1（如果尔普开枪得到的收益是 2）。相反，如果陌生人是普通的牛仔，那么尔普采取僵持策略得到的收益为 6（如果开枪得到的收益为 5）。所以，尔普开枪是最优策略。于是，策略组合 C 是一个贝叶斯纳什均衡。

第 11 章

11.1 基于文章中的分析，我们只需要确定经理不雇用每周工作 60 小时的实习生是最佳选择。现在，雇用该实习生获得的期望收益是 $p\times25+(1-p)\times100$，则一定有：

$60\geqslant p\times25+(1-p)\times100$ 或 $p\geqslant40/75$

只要这个工作 60 小时的实习生为懒惰型的概率至少是 40/75，此策略组合即是一个精炼贝叶斯纳什均衡。

11.2 思考这个懒惰型的实习生。工作 40 小时且未被雇用，可收益 20；而工作 60 小时并被雇用，其收益为 55。因此，她宁愿工作 60 小时，在这种情况下，该实习生的策略并非最优。

11.3 思考下列策略组合及信念：卖家的策略包括在汽车质量为劣质时以 P' 出售汽车以及在汽车质量为中等或优质时不出售汽车。买家的策略包括当价格不超过 P' 时购买汽车，价格超过 P' 时则不购买。在

买家的信念中，不管价格是多少，赋予汽车为劣质的概率为 1。显然，这种信念是一致的。当且仅当 $P' \leqslant 12\,000$ 时，买家策略是最优的，在这种情况下，价格没有超过买家对劣质汽车的估值。当且仅当 $P' \geqslant 10\,000$ 时，有劣质汽车的卖家策略是最优的；当且仅当 $P' \leqslant 15\,000$ 时，有中等质量汽车的卖家选择惜售是最优的；而当且仅当 $P' \leqslant 20\,000$ 时，有优质汽车的卖家选择惜售为最优。综合考虑上述情况，我们发现 p' 一定满足 $10\,000 \leqslant P' \leqslant 12\,000$。

第 12 章

12.1 一个分离均衡是发送信息人（汤姆），当他的类型是友好的时，发送信息停止；当类型是自私的时，发送信息继续。信息接受者（杰西卡）对信息停止采取行动开始，对信息继续采取行动隐藏。当信息接收者看到信息停止，认为发送信息人类型为友好的的概率为 1，当信息接收者看到信息继续，认为发送信息人类型为自私的概率为 1。第二个半分离均衡是当汤姆的类型是友好的时，发送信息继续；当类型是自私的，发送信息停止。杰西卡对信息停止采取行动隐藏，对信息继续采取行动开始。当杰西卡看到信息为停止，认为信息发送人的类型是自私的概率为 1，当杰西卡看到信息继续，认为信息发送人类型是友好的的概率为 1. 而对于胡说均衡，考虑到她的先验信念，杰西卡的最优行动是开始，因为这使她得到的收益为 $0.4 \times 3 + 0.6 \times 1 = 1.8$，而采取行动隐藏得到的收益为 $0.4 \times 1 + 0.6 \times 2 = 1.6$。一个胡说均衡是不管是哪种类型，汤姆发送信息停止，且杰西卡采取行动开始。第二个均衡是不管是哪种类型，汤姆发送信息继续，且杰西卡采取行动开始。

12.2 一贯的信念使加里分配概率的情况为：当消息为立刻时，莱斯莉的类型为低的概率为 $1/3=(1/6)/[(1/6)+(1/3)]$，类型为中的概率为 2/3，当消息为稍后时，莱斯莉的类型为高的概率为 1。按照这种信念，如果加里看到消息立刻，他从行动 a 中得到的期望收益为 $(1/3) \times 2 + (2/3) \times 3 = (8/3)$，从行动 b 中得到的期望收益为 $(1/3) \times 1 + (2/3) \times 2 = (5/3)$——所以采取行动 a 更优。当看到消息稍后，他从行动 a 中得到的期望收益为 1，从行动 b 中得到的期望收益为 2，所以采取行动 b 更优。于是加里的策略满足序贯理性。现在让我们来看莱斯莉的策略。给定加里的策略，她知道如果她选择消息立刻，那么他会采取行动 a，如果她选择消息稍后，那么他采取行动 b。所以，如果她的类型为低，那么发送消息立刻使她得到的收益为 2，发送消息稍后得到的收益为 1；如果她的类型为中，那么发送消息立刻使她得到的收益为 4，发送消息稍后得到的收益为 3；如果她的类型为高，那么发送消息立刻使她得到的收益为 2，发送消息稍后得到的收益为 4。所以，莱斯莉的策略规定了每一种类型发送的最优消息。

第 13 章

13.1 如果 $\delta=0.8$，4 期收益流在第 t 期的现值 u_t 就是 $u_1+0.8 \times u_2+0.64 \times u_3+0.512 \times u_4$。因而，$A$ 收益流的现值为 44.28，B 收益流的现值为 45.96，C 收益流的现值为 41.16，B 最优。

13.2 尽管阶段博弈并非只有唯一的纳什均衡——（高，低）和（低，高）都是其纳什均衡，但仍要注意这二者导致了同样的收益，也就是说，每个参与者都得到了收益 3。这表示不管参与者在第二阶段做了什么，他们在最终阶段都会得到收益 3。此外，他们在第二阶段的遭遇是一个有效的一次性博弈，因为从第一阶段获得的收益是无法改变的（过去的就过去了），且无论在第二阶段发生什么第三阶段的收益都是 3。因此，第二阶段参与的子博弈精炼纳什均衡一定是该阶段博弈的纳什均衡。对第一阶段亦同理。一个三阶段博弈的子博弈精炼纳什均衡是阶段博弈纳什均衡的任意排序，例如，参与者 1 在第一阶段选择低，第二阶段选高，第三阶段选高；参与者 2 在第一阶段选高，第二阶段选低，第三阶段选低。

13.3 思考第一阶段或未来的某些阶段，在该阶段参与者从未选择过 y。假如一个参与者选择了 x，则她期望今天获得收益 5，并且在将来，该收益的现值为 $5/(1-\delta)$；如果她选择了 y，则她的当前收益为 7；如果参与者都选择了 z，则她期望未来收益流为 2。该收益的现值为 $7+\delta\,[2/(1-\delta)]$。最后，她还可以在今天选择 z，这会带来 3 个当前收益和 5 个未来收益。最后这个选择明显劣于选择 x，因

为这两种选择都会产生同样的未来收益流，但 x 的当前收益较高。因此，达到均衡所需要的是至少令选择 x 与选择 y 一样好：

$\frac{5}{1-\delta} \geqslant 7+\delta\left(\frac{2}{1-\delta}\right)$，简化不等式得到：$\delta \geqslant \frac{2}{5}$

现在考虑一段历史记录，在过去的某段时间里，一个参与者选择了 y。从现在起参与者预期选择 z，且由于这是一个阶段博弈纳什均衡，显然为最优解。

总而言之，当且仅当 $\delta \geqslant \frac{2}{5}$ 时，该策略组合为子博弈精炼纳什均衡。

13.4 令：

$V=100+\delta\times 50+\delta^2\times 25+\delta^3\times 100+\delta^4\times 50+\delta^5\times 25+\cdots$，

等式两边同乘以 δ^3：

$\delta^3 V=\delta^3\times 100+\delta^4\times 50+\delta^5\times 25+\delta^6\times 100+\delta^7\times 50+\delta^8\times 25+\cdots$

再用第一式减去第二式，求出 V：

$V-\delta^3 V=100+\delta\times 50+\delta^2\times 25$

$$V=\frac{100+\delta\times 50+\delta^2\times 25}{1-\delta^3}$$

第 14 章

14.1 往事可以分成三种类型。首先，假设这段历史是这样的，两家拍卖公司都将收费标准设为 8%，则均衡条件为：

$$\frac{5}{1-\delta} \geqslant 7+\delta\times 0+\delta^2\times 0+\delta^3\left(\frac{5}{1-\delta}\right)$$

现在思考另一种情形，在当前阶段中将会开始存在惩罚，则均衡条件为：

$$0+\delta\times 0+\delta^2\left(\frac{5}{1-\delta}\right) \geqslant 1+\delta\times 0+\delta^2\times 0+\delta^3\left(\frac{5}{1-\delta}\right)$$

最后一种情况是，如果两家拍卖公司处于惩罚的第二阶段，则均衡条件为：

$$0+\delta\left(\frac{5}{1-\delta}\right) \geqslant 1+\delta\times 0+\delta^2\times 0+\delta^3\left(\frac{5}{1-\delta}\right)$$

14.2 假设在第一阶段或在过去所有的奇数阶段，结果都是（a，z），且在过去所有偶数阶段结果都是（c，x）。如果当前处于奇数阶段，参与者 1 倾向于选择 a。当且仅当：

$$1+\delta\times 12+\delta^2\times 1+\delta^3\times 12+\cdots \geqslant 3+\delta\left(\frac{4}{1-\delta}\right)$$

或 $\frac{1+12\delta}{1-\delta^2} \geqslant 3+\delta\left(\frac{4}{1-\delta}\right)$

时这一策略为最优策略。如果当前处于偶数阶段，参与者 1 倾向于选择 c，显然这是最优策略，因为该策略使她的当期收益与未来收益都得到最大化。对其他任一历史记录来说，假设参与者 1 选择 b，而这无疑也是最优策略，因为参与者 2 被假定为在当前和未来所有阶段都选择 y。根据对称性，同等条件适用于参与者 2。

14.3 当 $\frac{10}{1-\delta} \geqslant 12+\delta\left[0.3\times\left(\frac{3}{1-\delta}\right)+0.7\times\left(\frac{10}{1-\delta}\right)\right]$

$\Rightarrow \delta \geqslant \frac{2}{4.1} \approx 0.49$ 时，没有 ABM 系统优于较低等级的 ABM 系统；

当 $\frac{10}{1-\delta} \geqslant 18+\delta\left[0.75\times\left(\frac{3}{1-\delta}\right)+0.25\times\left(\frac{10}{1-\delta}\right)\right]$

$\Rightarrow \delta \geqslant \frac{8}{13.25} \approx 0.60$ 时，没有 ABM 系统优于较高等级的 ABM 系统。

这对策略是一个折现系数至少为 0.6 的子博弈精炼纳什均衡，在较弱的监控技术条件下，折现系数至少要达到 0.74。

第 15 章

15.1 在一个人的生命的倒数第二个阶段，当且仅当

$\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta\times 6\left(\frac{T-1}{T}\right)N \geqslant \left[6\left(\frac{T-1}{T}\right)N-1\right]$ 时，

她发现最优的策略是采取合作的态度。

在她生命的 $T-2$ 个阶段，均衡条件是：

$$\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta\times\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta^2\times 6\left(\frac{T-1}{T}\right)N \geqslant 6\left[\left(\frac{T-1}{T}\right)N-1\right],$$

更具体地来说，在 t 阶段中均衡条件为：

$$\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta\times\left[6\left(\frac{T-1}{T}\right)N-10\right]+\cdots+\delta^{T-t-1}\times\left[6\left(\frac{T-1}{T}\right)N-10\right]+\delta^{T-t}\times 6\left(\frac{T-1}{T}\right)N \geqslant 6\left[\left(\frac{T-1}{T}\right)N-1\right]$$

15.2 若卖家网页没有任何负面评价，买家的最优策略是购买；若卖家网页有一条或多于一条负面评价时，买家的最优策略是不买。在考虑卖家的最优策略时，最关键的是当卖家网页没有任何负面评价的时候。规定的行动是以 20 的价格提供质优产品，最好的替代行动是以 20 的价格提供劣质产品（和这章的逻辑相同）。当：

$$\frac{7}{1-\delta} \geqslant 18+\delta\left(\frac{0}{1-\delta}\right) \text{或} \delta \geqslant \frac{11}{18}$$

时，规定的策略为更优策略。所以，当且仅当 $\delta \geqslant 11/18$ 时，这个策略组合是一个子博弈精炼纳什均衡。

第 16 章

16.1 当 $V>C$ 时，鹰策略（就是说，$p=1$，所以鹰策略被选择的概率为 1）是一个演化稳定策略。这个策略为强演化稳定策略的条件为：

$F(1,1)>F(q,1)$，$0 \leqslant q<1$，

或者：

$$\left(\frac{V-C}{2}\right)>q\times\left(\frac{V-C}{2}\right)+(1-q)\times 0,\ 0\leqslant q<1$$

等同于 $1>q$，$0 \leqslant q<1$。

最后一个条件明显是成立的。当你看了图 16—1 的收益矩阵，你会明显发现鹰策略是一个演化稳定均衡。鹰策略严格优于鸽策略，意思就是当和自己采取相同策略的参与者博弈时，采取鹰策略得到的收益高于采取鸽策略得到的收益；当和采取鸽策略的参与者博弈时，得到的收益也更高。

16.2 存在两个严格纳什均衡：（慢，慢）和（快，快）。两个都是对称性的，所以两个都是（强）演化稳定策略。另一个可能成为演化稳定策略的情况是在混合策略的情况下。所以，让我们在混合策略中找到一个纳什均衡。如果 p 是选择慢的概率，那么当且仅当：

$$p\times 6+(1-p)\times 0=p\times 2+(1-p)\times 1 \Rightarrow p=\frac{1}{5}$$

$p=1/5$ 不是一个强演化稳定策略，要成为弱演化稳定策略，就必须满足以下条件：

（1）$F\left(\frac{1}{5},\frac{1}{5}\right)=F(q,\frac{1}{5})$，$q\neq\frac{1}{5}$；

（2）$F\left(\frac{1}{5},q\right)=F(q,q)$，$q\neq\frac{1}{5}$。

因为对所以的 q 来说（包括 $q=\frac{1}{5}$），$F\left(q,\frac{1}{5}\right)=\frac{6}{5}$，满足条件（1）。条件（2）是：

$$\left(\frac{1}{5}\right)\times[q\times 6+(1-q)\times 0]+\left(\frac{4}{5}\right)\times[q\times 2+(1-q)\times 1]>q\times[q\times 6+(1-q)\times 0]+(1-q)\times[q\times 2+(1-q)\times 1],$$

$q\neq\frac{1}{5}$

经过变形，我们得到：

$$\frac{10q+4}{5}>1+5q^2$$

所以不能满足这个条件。比如，如果 $q=1$，那么不等式的左边等于 $\frac{14}{5}$，就不满足 $\frac{14}{5}>6$。所以，在混合策略中不存在演化稳定策略。那么这场博弈中有两个演化稳定策略：每个人都选择慢，或者每个人都选择快。

16.3 使它成为演化稳定策略的条件是：

$\left(\frac{2}{3}\right)\times F$（强势，强势）$+\left(\frac{1}{3}\right)\times$ F（强势，温和）$>$ F(温和，强势)，

或者：$\left(\frac{2}{3}\right)\times 1+\left(\frac{1}{3}\right)\times 8>3$；

或者：$\frac{10}{3}>3$。

第 17 章

17.1 让 p^t 代表在 t 代中被赋予策略跳的比例，我们得到：

采取跳得到的适应值为：$p^t\times 3+(1-p^t)\times 2=2+p^t$，

采取跑得到的适应值为：$p^t\times 1+(1-p^t)\times 2=2-p^t$。

所以：

$p^{t+1}>p^t$，$2+p^t>2-p^t$ 且 $p^t>0$

所以，吸引子就是使所有的成员被赋予策略跳。

17.2 让 b^t 代表在 t 代群体成员中被赋予策略黑的比例，所以，$1-b^t$ 就代表被赋予策略白的比例。采取策略黑得到的适应值为：

F^t（黑）$=b^t\times 5+(1-b^t)\times 1=1+4b^t$

采取白策略得到的适应值为：

F^t（白）$=b^t\times 1+(1-b^t)\times 2=2-b^t$

由复制者动态，我们知道当且仅当：

F^t（黑）$>F^t$（白）或者 $1+4b^t>2-b^t$ 或者 $b^t>1/5$ 时，采取黑策略的比例会上升；

而 F^t（黑）$<F^t$（白）或者 $1+4b^t<2-b^t$ 或者 $b^t<1/5$ 时，采取黑策略的比例会下降。

总之，我们得到：

如果 $b^t>1/5$，那么 $b^{t+1}>b^t$；

如果 $b^t=1/5$，那么 $b^{t+1}=b^t$；

如果 $b^t<1/5$，那么 $b^{t+1}<b^t$。

于是，就存在两个吸引子。每个成员采取黑策略是局部稳定的；只要当前被赋予黑策略的比例大于 1/5，群体结构就会逐渐收敛到那个静止点。每个人

采取白策略也是局部稳定的；只要当前被赋予黑策略的比例小于 1/5，群体结构也会逐渐收敛到那个静止点。群体中 20% 的成员采取黑策略也是一个静止点，但是这种状态下不是局部稳定状态。如果群体中采取黑策略的成员小于 20%，群体结构就会演化为没有成员被赋予黑策略（而不是 20% 被赋予黑策略），而如果群体中采取黑策略的成员大于 20%，那么群体演化为所有的成员都被赋予黑策略。

术语表

Action 行动：扩展式博弈中参与者在决策节点所做的选择。

Attractor 吸引子：它是这样一个静止点，即群体结构一旦接近这个静止点，群体结构最终会停留在这个静止点。

Babbling equilibrium 胡说均衡：廉价谈话博弈中的混同均衡。

Backward induction 逆向归纳法：子博弈精炼纳什均衡解的扩展式博弈求解方法。最后一个子博弈给出为纳什均衡，其他子博弈替代为与均衡参与相关联的收益。该过程一直重复到全部博弈求解出。

Basin of attraction 吸引域：能使群体结构最终达到一个吸引子的一系列群体结构比例集合。

Bayesian game 贝叶斯博弈：标准博弈的一种变体，在贝叶斯博弈中，将"自然"先行引入被赋予私人信息的参与者中。

Bayes-Nash(Bayesian)equilibrium 贝叶斯纳什(贝叶斯)均衡：贝叶斯博弈的一个解决概念。贝叶斯纳什（贝叶斯）均衡是一个策略组合，这个策略组合规定了每个参与者每种类型的最优行动，给定其他参与者的策略和关于其他参与者类型的信念。

Bayes's rule 贝叶斯法则：利用获取的新信息更新信念的方法。令 x 为随机变量，Prob（x）代表在 x 上的概率分布。例如，Prob（x'）表示的是 $x=x'$ 的概率。通过认识到另一随机变量 y 的存在从而获得新的信息。设 Prob（x，y）表示 x 和 y 的联合概率。根据观察到的 $y=y'$，对 x 的新信念用 Prob（$x|y'$）表示。按照贝叶斯法则，这些基于条件的信念可以由以下公式表示，即 $\text{Prob}(x|y')=\dfrac{\text{Prob}(x,\ y')}{\text{Prob}(y')}$。

Best reply 最优反应：一个参与者根据对其他参与者所选策略的信念采取最大化自身收益的策略。

Branch 分支：扩展式博弈的一个部分，代表了参与者的某一项行动。

Cardinal payoffs 基数效用：描述了参与者偏好强度的收益，而不仅仅是多种选择如何排列的问题。

Cheap talk game 廉价谈话博弈：在一场可以表明意图的博弈中，发送信息人采取的不直接影响参与者支付收益的零成本行动，这个模型就是廉价谈话博弈。

Commitment 承诺：一个参与者限制自己未来选择的行为。一个参与者通过限制他 / 她未来的选择或改变他 / 她未来的收益做出的承诺。

Common knowledge 共同知识：就是所有的参与者在无穷递归意义上均知悉的事实。比如，如果（1）杰克和吉尔都知道昨天下雨这个事实；（2）杰克知道吉尔知道昨天下雨了，吉尔也知道杰克知道昨天下雨了；（3）杰克知道吉尔知道杰克知道昨天下雨了，吉尔也知道杰克知道吉尔知道昨天下雨了；等等，对杰克和吉尔来说，昨天下雨这个事实就是共同知识。

Common prior assumption 共同优先假设：贝叶斯博弈的一个特性，指在参与者类型决策之前，每一个参与者对所有参与者的类型抱有同样的信念，且这些信念在参与者中是共同知识。

Common value 共同价值：在拍卖中，所有的参与者对被拍卖的物品赋予相同的价值，尽管他们并不确定物品的真正价值是多少。

Complete preference 完全偏好：若一个人能够区分任意两个选项并说出首选哪个或是说出它们是否完全相同，则称偏好是完全的。

Conflict 冲突：博弈的一种特性，表现为比较策略组合时，如果一个参与者的收益更高，那么另一参与者的收益就更低。

Congestion 挤出效应：收益函数的一个特性，即当更多的参与者使用某一策略时，也采用这一策略的相对吸引力更小的情况。

Consistent beliefs 一致信念：在观察到某一参与者行动的条件下，对这一参与者类型的信念需具有一致性。当参与者更新对另一参与者类型的信念时，必须考虑到哪一类型的参与者才会采取这样的行动。要做到这一点，需要对参与者的策略做出猜测，然后判断每一类型的参与者会采取观察到行动的概率。

Constant-sum game 常和博弈：在博弈中，参与者从所有策略组合中得到的收益总和等于同一个数。这些博弈属于纯冲突博弈类型。

Continuous strategy set 连续策略集：由实数的区间组成的一个策略集合。

Cooperation 合作性均衡：在重复博弈中，当参与者选择的行动可导致所有参与者都获得更高的平均收益，而非达到策略博弈的纳什均衡时，该均衡就显示为合作性均衡。

Coordination game 协调博弈：所有参与者选择相同策略时使收益最大化的一种博弈类型。

Decision node 决策节点：在扩展型博弈中一个参与者选择一个行动的节点。

Decision tree 决策树：对参与者的行动顺序和可采取的行动进行的图形描述。

Discount factor 折现系数：用来推导一系列收益的现值采用的权重因子。较低的折现系数表示参与者未来收益的贴现越少。

Discount rate 折现率：如果折现系数为 δ，则折现率 r 满足 $\delta=1/(1+r)$。在某些情况下，折现率相当于利率。

Dominance solvable 占优可解的：对于每个参与者，只有一个策略存在于重复剔除严格劣势策略中的博弈特性。

Dominant strategy 占优策略：一个严格优于其他所有策略的策略。

Equilibrium payoff dominance 均衡博弈：纳什均衡的一个特性，表现为没有一个均衡中出现每个参与者有一个严格更高收益的情况。

Equilibrium play 均衡参与：对一个扩展式博弈来说，行为结果由参与者使用均衡策略引发。

Evolutionarily Stable Strategy 演化稳定策略（ESS）：是指如果占群体绝大多数的个体选择演化稳定策略，那么小的突变者群体就不可能侵入到这个群体，即采取演化稳定策略得到的适应值高于采取突变策略得到的适应值。

Evolutionarily Stable Strategy Profile 演化稳定策略组合（ESSP）：对一个多元群体的场景来说，如果占多元群体绝大多数的群体选择演化稳定策略组合，那么小的突变者群体就不可能侵入到这个群体，即采取演化稳定策略组合得到的适应值高于采取突变策略得到的适应值。

Evolutionary game theory 演化博弈理论：博弈理论在生物学上的应用，演化博弈理论这个概念是基于自然选择而非理性。

Exhaustive search 穷举搜索：通过核查每个策略组合是否满足成为博弈的解的条件（如纳什均衡条件）实现求解的方法（如多重纳什均衡的解）。

Expectation 期望：某一事件在同一情况下重复多次的平均值。

Expected payoff 期望收益：收益在同一情况下重复多次的平均值。是一个参与者收益的加权平均值，某一收益的权数是该收益能够收到的概率。

Expected value 期望值：参见期望。

Experiential learning 体验式学习：通过经验来学习。在对某人将表现如何进行决策的情况下，经验式学习意味着“根据过去的行为形成信念”。

Extensive form game 扩展式博弈：对战略形势的图形化描述，它描述了参与者在博弈过程中所遇到决策问题的序列结构、当他 / 她做决定时可供他 / 她选择的行动方案以及每一种博弈结果中参与者得到的收益值。

Finite horizon 有限次视野：参与者中遭遇者的人数有限时存在的一种策略情境。

Finitely repeated game 有限次重复博弈：具有有限视野的重复博弈。

First-price auction 首价拍卖：一种拍卖形式，当中所有的出价人同时竞价，最高出价的竞价人支付相应的竞拍价格后获得拍品。

Focal point 焦点：具有显著特性的一个策略组合。

Free riding 搭便车：个体不采取和群体中其他人一样的行动而从群体中获益。

Game of complete information 完全信息博弈：参与者拥有共同知识的博弈。

Game of incomplete information 不完全信息博弈：对参与者来讲非共同知识的博弈——比如，一个参与者的收益只有自己知道。

Game theory 博弈论：关于博弈的理论。

Grim-trigger stategy 冷酷策略：在这种策略下，合作参与中的任意偏差可引起一个阶段博弈纳什均衡的永久性逆转。

History 历史：过去参与行为的累积。t 期间的历史是由参与者在 $1,2,\cdots,t-1$ 期间所做的全部选择组成的。

Imperfect information 不完美信息：在扩展式博弈中，至少有一个信息集不是单节点信息集。换句话说，博弈中至少有一个参与者在博弈的某个阶段必须采取行动，但他 / 她并不知道博弈的状况。

Imperfect recall 不完美记忆：扩展式博弈中的一个概念，它指的是一个参与者在博弈过程中缺失了关于博弈的信息。比如，如果一个参与者忘记自己以前知道或做过的事情或如果一个参与者在博弈的某个阶段可以区分两个决策节点，但在下一个信息集中不能区分这两个决策节点，这个参与者满足"不完美记忆"的要求。

Incomplete information 不完全信息：非共同知识的博弈的一个特性。

Incredible threat 不可置信的威胁：对特定情况（或信息集）提出的一种行动方案，且如果这个特定情况发生，对这个参与者来说，这个行动方案不是他的最佳方案。

Indefinite horizon 不确定视野：参与者中遭遇者的人数随机时存在的一种策略情境。此外，参与者绝不相信当前遭遇将会是最后一个。由于不确定视野的存在，参与者通常不确定他们未来是否将会与其他每个人遭遇。

Indefinitely repeated game 不确定重复博弈：具有模糊视野的重复博弈。

Independence 独立性：随机事件的特征，一个事件结果的发生与另一事件结果发生的可能性无相关。

Independent private value 独立私人价值：在拍卖中，一个参与者对物品的估价是独立的，和其他参与者的估价高低并无联系。

Infinite horizon 无限次视野：参与者中遭遇者的人数无限时存在的一种策略情境。

Infinite regress 无穷后退：在该情境中，某个问题的解取决于求解一个关于该问题的无穷序列，其中每个问题的解取决于序列中之前的问题。换句话说，一个问题的解取决于第二个问题的解，第二个问题的解又取决于第三个问题，以此类推。

Infinitely repeated game 无限次重复博弈：具有无限次视野的重复博弈。

Information set 信息集：在一个博弈中，参与者无法区分的决策节点的集合。

Information cascade (Informational herding) 信息瀑布（信息羊群效应）：在该情境中，人们持续行动，且最初几人的行为暴露出来的信息掩盖了后来者的个人信息，以至于他们的行为独立于他们的私人信息之外。

Initial node 首节点：扩展式博弈中的一个决策节点，在这个决策节点，博弈中的参与者所做的第一个选择。

Iterative deletion of strictly dominated strategies (IDSDS) 重复剔除严格劣势策略：一个剔除所有严格劣势策略的过程，这一过程会多次重复直到博弈中没有策略可以被剔除。这一过程之后所剩的策略被称之为"存在于重复剔除严格劣势策略中"的策略。

Locally stable rest point 局部稳定静止点：参照吸引子。

Lying 说谎：带有欺骗意图而做不真实的陈述。说谎的人通常会发现他们的肢体外观有着令人眼花缭乱的表现。

Maximin solution 极大极小解：每个参与者能在其中运用其极大极小策略的策略组合。

Maximin strategy 极大极小策略：参与者将其期望收益最大化的策略，即假定其他参与者根据该参与者的策略选择能使其期望收益最小的策略。极大极小策略将参与者的最小期望收益最大化。

Message 消息：一个参与者采取的零成本行动，意在传递信息或意图。

Mixed strategy 混合策略：参与者的纯策略集上的概率分布。

Mutation 突变：在演化博弈理论中，群体成员采取策略的随机改变。

Mutual interest 共同利益：博弈的一种特性，表现为比较策略组合时，如果一个参与者的收益更高，则其他参与者亦随之更高的情形。

Nash equilibrium 纳什均衡：一个每个参与者针对其他参与者所选的策略选择最大化自身收益的策略组合。

Nature 自然：一个在博弈中其行为由外界指定的参与者。自然是对随机行为的典型假设——其行为由基于各种行为之上的概率总结出来——而且意在表现参与者在其环境中关于随机的信念。

Network effect 网络效应：收益函数的一个特性，表现为更多的参与者选择某一策略时，这一策略的价值就更大的情形。

Non-constant-sum game 非常和博弈：在此博弈中，参与者从所有策略组合中得到的收益总和不等于同一个数。该博弈尽管能继续保持冲突，但还是可以为共同利益提供基础。

Normal-form game 标准式博弈：策略式博弈的另一种叫法。当策略式博弈被错误指定时，它指的是一个“阿比标准式博弈。”（参见《新科学怪人》，*Young Frankenstein*）

n-tuple 多元组：多物体的集合。例如，一个 *n* 人博弈的策略组合就是一个策略多元组，每个参与者一个。

Optimal strategy 最优策略：一个使参与者收益最大化的策略。

Ordinal payoffs 序数效用：描述参与者如何对不同选择进行排序的收益。

Outguessing game 猜谜博弈：一种博弈类型，其中参与者的最优反应是不使用使其他参与者所选策略成为最优策略的策略。换句话说，参与者所选的策略超出了其他参与者的预期。

Overlapping generations 世代交替：在群体中，由不同代际的成员组成。在任何时候，群体中都有即将离去的人和新进的成员。

Partially optimal strategy 部分最优策略：在重复博弈中，如果每个时期每段历史中，该策略规定的行为与选择其他任意行为相比，可带来最高收益，并假设在未来期间参与者的行为由策略来决定，则该策略是部分最优的（有子博弈精炼纳什均衡的感觉）。

Payoff 收益：一种用来衡量与博弈某一特定结果相关的参与者盈利状况的标准。

Payoff dominance 收益占优：策略组合的一个特性，表现为其他策略组合中不存在使每个参与者有一个严格更高收益的情况。

Payoff-dominant Nash equilibrium 收益占优纳什均衡：与其他的纳什均衡相比，一个可以使每个参与者得到更高收益的纳什均衡。

Payoff-dominated Nash equilibrium 收益不占优纳什均衡：一个当中出现另一个使每个参与者的收益更高的纳什均衡的纳什均衡。

Payoff function 收益函数：一个参与者对每个策略组合所赋予的一个收益。

Perfect Bayes-Nash equilibrium 精炼贝叶斯纳什均衡：由一个策略组合和后验集合组成。鉴于对其他参与者类型的信念，每个参与者的策略对每个信息集合都是最优反应，同时这些后验总是与参与者的策略和贝叶斯法则相一致。

Perfect information 完美信息：在扩展式博弈中，所有的信息集都是单节点信息集。也就是说，当一个参与者采取行动时，他 / 她总是知道博弈中所有的情形。

Perfect recall 完美记忆：扩展式博弈中的一个概念，它指的是在博弈中，参与者永远记得他们知道的信息。（进一步解释可以参照“不完美记忆”的概念，“不完美记忆”描述的是参与者没有“完美记忆”的情况。）

Phase diagram 相图：用来描述群体混合策略随着时间如何演化的轨迹。

Pooling equilibrium 混同均衡：在不完全信息博弈中，存在着一个其中每个拥有私人信息的参与者对每种类型采取相同行动的均衡。

Pooling strategy 混同策略：参与者对每种类型采取相同行动的策略。

Posterior beliefs 后验信念集合：参与者根据获得的新信息更新先验信念后的信念。后验信念是通过使用贝叶斯法则而产生的。

Posterior belief 后验信念：你耐着性子听完一个两个小时的课程后渴望的事情。

Preemption game 抢先博弈：在这一博弈中，每位参与

者都在决策何时采取一个特定行为，当满足（1）在其他参与者之前采取行动，且（2）在采取行动前多等待一会儿时，该参与者的收益达到更多。

Preplay communication 前博弈沟通：扩展式博弈的最初组成部分，针对优先选择免费信息而不是付费信息的参与者。

Present value 现值：一组收益流的现在价值，是参与者的与其他收益流无关的当期收益。如果这个基于 $1,2,\cdots,T$ 期间的单个期间收益流用 $v_1,v_2,v_3,\cdots,v_T$ 表示，则该收益流的现值为 $v_1+\delta v_2+\delta_2 v_3+\delta_3 v_4+\cdots+\delta^{T-1}v_T$，其中 δ 是折现系数。一组收益流的现值也被称为单个期间收益折现之和。

Prior beliefs 先验信念：参与者对某一未知事物最初的信念。

Prisoners' Dilemma 囚徒困境：这一表示所属的撇号应该放在单数之后还是复数之后呢？也就是说，究竟是囚徒困境还是囚徒们的困境？博弈理论家对这一问题的辩论仍在如火如荼地进行。

Probability 概率：无限重复某一情形时，发生一个事件的相对频率。它的值在 0 和 1 之间。

Probability distribution 概率分布：某一随机事件发生概率的集合，即每一个可能发生这一事件的概率。这些概率的总和为 1。

Projection 映射：一个策略组合在子博弈中的映射，是为该子博弈中的信息集规定行为的策略组合的组成部分。也称为子策略组合。

Pure strategy 纯策略：参与者的不包含随机选择的策略。

Quid pro quo 交换：一个人在交换中提供事物交换另外事物的策略情境。是拉丁文中的“以物易物”,在《王牌大间谍》中被误用为“squid pro row”。

Random 随机：不可预料的。

Random event 随机事件：不可预料的事件。

Random variable 随机变量 其（数）值不可预知的数量。

Randomized version of a game 博弈的随机版本：参与者的纯策略集被混合策略集所取代的一种博弈，是原始博弈纯策略集的概率分布。在对混合策略组合进行评价时，参与者收益用期望收益来代替。

Rational 理性：基于对环境和其他参与者者行为的信念采取最大化自身收益的行动。

Rationality is common knowledge 理性是共同知识：一种无限的信念层次，其中参与者 1 知道参与者 2 是理性的，参与者 3 知道参与者 1 知道参与者 2 是理性的，以此类推。

Rationalizable 可理性化的：如果一个策略与理性是共同知识相一致，那么这一策略就是可理性化的。也就是说，基于与理性是共同知识相一致的信念，一个参与者选择的策略就是最优策略。

Rationalizable strategy 可理性化策略：根据其他参与者所选策略与理性是共同知识相一致的信念，一个参与者选择的最优策略。

Reciprocal altruism 互利主义：是指一个人（或其他生物）从事一种代价较高的活动，该活动对其他人有益并期望在未来得到回报的一种策略情境。

Regular subtree 标准子树：包含所有信息集的子树，在子树中至少有一个节点。

Repeated game 重复博弈：阶段博弈的重复。

Replicator dynamic 复制者动态：具体说明群体中采取一种策略的比例及得到的适应值和群体平均适应值的关系。更具体地来说，当采取一种策略得到的适应值大于群体得到的平均适应值，那么采取这种策略的比例就增加，反之，减少。

Rest point 静止点：群体结构达到的一个点且停留在那个点，这个点就是静止点。

Second price auction 二级价格拍卖：一种拍卖方式，其中所有的出价人同时出价，出价最高的竞价人获得拍品，而他或她只需支付第二最高出价的价格。

Self-awareness 自我认知：有智慧生物的一种特性，是对自我存在的感知。

Self-enforcing agreement 自我实施协议：参与者之间基于个人利益而实施彼此之间达成的一个协议，如纳什均衡就是一个自我实施的协议。

Semipooling equilibrium 半混同均衡：一个当中出现有私人信息的参与者使用半混同策略的纳什均衡。

Semipooling strategy 半混同策略：在贝叶斯博弈中，一个既不是混同策略也不是分离策略的策略。采用这样一种策略，一个参与者的行动只对他或她的类型提供了一些信息，甚至在某些情况下可能会提供全部信息。这一策略也称为半分离策略。

Semiseparating equilibrium 半分离均衡：一个当中出现有私人信息的参与者选择半混同策略的纳什均衡。

Semiseparating strategy 半分离策略：参见半混同策略。

Separating strategy 分离策略：一个当中出现每个参与者对每种类型的发出者选择不同行动的策略。

Separating equilibrium 分离均衡：在贝叶斯博弈中，每个拥有私人信息的参与者对每种类型的发出者采取不同行动的一种纳什均衡。

Sequential rationality 序贯理性：作为一种要求，表现为在博弈的每一阶段，参与者根据对其他参与者行动的信念采取一个策略，这一策略指定一个最优行动。

Set of players 参与者集：博弈中所有采取行动的个体的集合。

Signaling game 信号传递博弈：一种不完全信息下的两人博弈，其中一个参与者（称为发出者）了解自己的类型后选择一种行动，而另一个参与者（称为接收者）根据观察到的这一行动确定自己的行动。

Simulated introspection 模拟内省：通过这种方法，另一个参与者的行为可以通过模拟其做决定的推理过程而被预知。

Simultaneous-move game 同时行动博弈：所有参与者在不知道其他参与者行动的情况下采取行动的一种扩展式博弈。

Sneetch 史尼奇：苏斯博士（儿童文学家、教育学家，译者注）虚构的一个人物，他生活中首要关心的事是决定要不要在肚子上放一个星星。

Social norm 社会规范：社会强制实施的行为标准。

Sociological Department 社会学部门：最初于 1914 年由福特汽车公司创建，其角色定位为为确保雇员的“集体行为”而对其行为进行监控并提出建议。是罗杰斯先生（Mr.Rogers，著名心理学家）与老大哥（Big Brother）的融合。

Stage game 阶段博弈：指重复博弈中被重复的那个子博弈。阶段博弈是重复博弈的构成要件。也被称为“单次博弈”。

Stage-game Nash equilibrium 阶段博弈纳什均衡：指阶段博弈的纳什均衡。

Strategic form game 策略式博弈：对策略形势的一种描述方法，包括 3 要素：（1）参与者集合，（2）每个参与者的策略集，和（3）每个参与者的支付函数。

Strategic interdependence 策略性相互依赖：两个或更多人之间遭遇的一个特性，对某人来说什么是最优的取决于其他人计划做什么。

Strategy 策略：策略是参与者如何对其他参与者的行动作出反应的行动规则。在扩展式博弈中，一个参与者的策略赋予他 / 她的每个信息集一种行动方案。

Strategy profile 策略组合：策略的集合。在博弈中，每个参与者有一个策略组合。

Strategy set 策略集：对一个参与者来说，所有可行的策略的集合。

Strict Nash equilibrium 严格纳什均衡：在一个纳什均衡中，每个参与者采取的策略都是最优策略。也就是说，采取其他任何一种策略得到的收益都更低。

Strictly dominated strategy 严格劣势策略：对其他参与者的每个策略组合都有另一个可以产生更高收益的策略，那么这个策略就是严格劣势策略。

Subgame 子博弈：含相关收益的标准子树。

Subgame perfect Nash equilibrium 子博弈精炼纳什均衡：对每个子博弈来说，如果它的子策略是纳什均衡，那么这个策略组合式子博弈精炼纳什均衡。

Substrategy (substrategy profile) 子策略（子策略组合）：对子博弈来讲，是指为该子博弈信息集规范行为的策略中的组成部分。一个子策略组合通常也被认为是一个映射。

Subtree 子树：一个非终端节点和树的所有其他节点构成的树型结构。

Summary statistic 概括统计量：对数字集合总结后的一个单个数字。例如，该收益流的现值就是一个概括统计量。

Symmetric game 对称博弈：如果（1）所有参与者的策略集合相同；（2）参与者选择相同策略得到的收益相同；（3）对调两个参与者的策略后他们的收益也被对调，那么这一博弈就是对称博弈。

Terminal node 终端节点：扩展式博弈中的最后一个节点。它与博弈中的序列行动相关。

Terminator 终止子：参与者在博弈中采取最后一个行动，随之达到终端节点。它应该是但还并不是一个真正的博弈论术语。

Theory of games 博弈理论：博弈论的另一表达方式。你真的以为我打算为此提供一个简要说明，好让你躲过本书不得不读的其余部分？

Theory-of-mind mechanism (ToMM) 心智理论机制：理

解他人拥有不同于我们自己的信仰和偏好的能力。

Tipping 倾向效应：收益函数的一种特性，表现为更多的参与者使用某一策略时，采用这一策略的相对吸引力更大的情况。注意不要与“轻拍母牛”以及“旁敲学生”相混淆，前者是指醉酒的大学生对毫无戒备的母牛表现出的行为，后者是指老师对上课睡觉的学生采取的措施。

Tit for tat 以牙还牙：指一个参与者由合作性行为开始，然后在另一个期间，去做另一个参与者在前一期间所做的事的策略情境。这一情境体现了“出来混总是要还的”这条准则，并且应该与更加宽厚的“己所不欲勿施于人”区别开来。

Tragedy of the commons 公地悲剧：在使用某一资源时，个人与所在集体之间发生的利益冲突。每个人都以自己的利益最大化为行动标准导致资源的过度开采与使用。

Transitive preferences 可传递性偏好：假定偏好是可以传递的，不论何时选项 A 都优先于选项 B 而选项 B 又优先于选项 C，随之，选项 A 优先于选项 C。

Trawling 搜罗法：随机搜索一个均衡。在对“此博弈的纳什均衡是什么”回答错误之后，学生会用搜罗法随机猜测一个答案，还做出一副他（她）知道他（她）在做什么的表情。不过一点用也没有。

Type 类型：在贝叶斯博弈中，一个参与者的类型就是自然赋予他（或她）的私人信息。例如，一个参与者的收益可能是私人信息时，他的收益就是自己所属的类型。

Type space 类型空间：贝叶斯博弈中，一个参与者可行类型的集合。

Undominated Nash equilibrium 非劣纳什均衡：当中不出现参与者使用弱劣策略的纳什均衡。

Utility 效用：对一个方案赋予一个数值，用来衡量一个参与者选择这个方案的“得失”状况。

Utility function 效用函数：所有可行的方案集合以及采取每个方案时得到的效用值。

War of attrition 消耗战：其中每个参与者都在决策何时采取一个特定行为的博弈，当（1）其他参与者在前述参与者之前采取同样行为，且（2）前述参与者更早地采取行动时，该参与者的收益会增多。

Weakly dominant strategy 弱占优策略：一个弱优于其他所有策略的策略。

Weakly dominated strategy 弱劣策略：如果对于其他参与者的每个策略，另一策略产生的收益至少等于这一策略的收益；对于其他参与者的某一策略，另一策略产生的收益严格大于这一策略的收益，那么这一策略就是弱劣策略。

Winner's curse 赢者的诅咒：当赢得拍卖品可能是坏消息时，赢者的诅咒现象发生。当投标者出价高于物品的价值时，即赢标者付出的代价太大，赢者的诅咒就发生了。

Zero-sum game 零和博弈：对于所有的策略组合，参与者的收益总和为零的博弈类型。或者说，一个常和博弈中收益总和为零的情形。



译者后记

“博弈”二字从字面上看起来就充满了古意，不禁会使人想到举棋对弈的场景，并让人感受到其中充满了智慧与深奥。实际上，从博弈的英文原词“Game”来看，博弈理论就是“游戏理论”，种种场景莫不像是一个个的小游戏、大游戏，是对各个领域的抽象与概括，我们每个人都可以运用“玩游戏”中获得的博弈知识在对弈中争得上风，成为“游戏”的赢家。

本书作者小约瑟夫・E・哈林顿是美国约翰霍普金斯大学的经济学教授。他曾任职于多家杂志编辑部；在不同学科的顶级杂志上发表过研究成果；还是备受美国各大名校推崇的经济学教科书——《反垄断与管制经济学》（*Economics of Regulation and Antitrust*）的作者之一。

正如作者书中所述，博弈理论本身不仅仅在经济学、商业、政治学、国际关系等领域有着重要应用，甚至在历史、文学、体育、犯罪、宗教、军事战略、计算机科学、生物学等领域都有着极其广泛的体现。“博弈论是一种适用于所有人类（甚至包括动物王国的某些成员）、会和所有人都不期而遇的推理工具”。事实上，从博弈论这门课程在国内高等教育中仅局限于面向某些传统学科开始，发展到今天，其适用对象已包括了绝大多数学生在内，它的确已经在普通高等教育中占有了重要的一席之位。

哈林顿教授是一位有着十五载（到本书出版之时，显然已不止这么久）博弈论教学经验的导师，不仅在博弈论研究方面有着深厚造诣，其在教学方面的重要心得更是对撰写本书大有裨益。他深刻地了解到什么会使学生困惑不已而无从下手，什么又会使学生茅塞顿开而恍然大悟；他知道如何避免让学生听课听到仿佛中了“昏睡咒”，也知道什么样的案例会让学生兴奋而专注。正是这些经验使得本书看起来堪称有趣。

本书采用直接且平易近人的风格，从案例入手使学生接触到相关概念，再通过一系列案例固化、增强对概念的理解。学习者将会分别在真实的、历史的和虚构的世界中运用博弈理论来搜寻策略情境，并加以分析。本教材案例范围颇广，涉及诸多主题，如伯罗奔尼撒战争、制药公司间的专利竞争、电视剧《幸存者》、会计丑闻、约会困境等。待到亲眼目睹学生在完成案例项目时体现出对博弈论的兴趣之广、热爱之深时，你会欣慰地发现这本教材的价值所在。

但特别要提到的是，尽管本教材中所运用的数学知识相对简单，但内容绝非仅仅停留在《博弈论傻瓜手册》或是《策略的完全傻瓜指南》这一层面，而是充满了大量对不完美信息博弈、不完全信息下的信号传递博弈（包括廉价谈话博弈）和远非简单的重复博弈的处理。随着结构的逐渐搭建，还引入了更为复杂的条件设置与概念。

同时，也正是基于多年的执教经验，哈林顿教授从怎样使学习效果更佳的角度出发，煞费苦心地在文中设置了“五保险”，即：一是在重要概念后附有理解测试；二是把定义与点拨用简洁、鲜明的图标凸显出来；三是将具有挑战性的谜题列示出来；四是在每个章节后都对本章要点进行了总结；五是在课后设置了练习题用以巩固。在翻译过程中，译者也被哈林顿教授这种又风趣又细心的风格深深打动。

全书内容分为六个部分。第一部分是概述，对博弈论进行了简要介绍；第二部分是对策略式博弈的求解，使

学生从简单概念转向复杂概念前进了一步；第三部分处理的是扩展式博弈问题；第四部分主要包括不完全信息博弈这一博弈论课程中最具挑战性的课题；第五部分致力于探讨重复博弈问题；本书以第六部分的演化博弈作为结束。每一部分都是由最基本的理论问题分析起步，再逐渐进入各个具体领域，并辅以案例进行讨论。各个部分既相互独立，又相互联系，教师的教学和学生的学习都有很大的自主性，因而使得本书不仅适合普通高校学生使用，企业和政府部门的管理者也都可以从中得到借鉴。

文中涉及的诸多术语在忠于作者原意的前提下，均采用了国内译著中的通行译法，如果某个术语有两种或两种以上译法的，一般采取使用最广泛的说法，以避免学习者在参阅其他书籍、文献时发生混淆。

虽然我学习经济学、管理学已多年，但也一样对“玄妙、深奥的大部头”心存畏惧，好在翻译本书的过程中完全没有“啃”的感觉，感受到的自始至终都是兴趣盎然。跟随着全书的脉络，我又将博弈论导论方面的知识重新温习了一遍，受益匪浅。在此，向湛庐文化的编辑致敬，是他们为我提供了这样好的学习机会，也为广大博弈论学习者提供了一部优秀的教材。相信对本书的学习者来说，学到的不仅仅是解开谜局的聪明答案，更多的是获得真正洞察人类行为的能力。

本书是由我和我的爱人李强老师共同翻译，并由我总纂定稿的。翻译工作耗时半年，又恰处本人博士论文撰写期间，非常感谢我的父母对我的大力支持，是他们悉心照顾着女儿的衣食起居，才使我们毫无顾虑地、全身心地投入到翻译工作中。特别要致谢的是我的两位导师：东北财经大学刘永泽教授和山西财经大学郭泽光校长，是他们严谨的态度和宽厚的胸怀，时刻给我以鼓励和鞭策，在此表示衷心感谢。

在翻译过程中得到了我们的老师、同事及几位硕士研究生的鼎力相助。山西财经大学的张遂老师、李艳苹老师、彭月兰老师、李端生老师、田祥宇老师、赵学文老师、郭宗文老师、李蕊爱老师以及上海大学的许金叶老师，为本书的翻译工作提供了许多有益的建议和指导，感谢他们一贯给予我俩的巨大鼓励。此外，我的亲友韩坚仪先生、吉晓岚女士、妹妹吉星、好友张栋、李莹、张敏、葛斌、硕士研究生李姣、白丽、仇月普及常晓燕，他们各自在专业知识、后勤保障上为我们提供了巨大帮助，还在图表整理、文字录入等环节做了大量工作，在此一并表示感谢。

最后，还要感谢湛庐文化的编辑，在时间紧、任务重的情况下得到了他们的理解和支持，才使得我们顺利完成翻译工作，在此对他们的辛劳表示衷心感谢！

由于才疏学浅，翻译中难免有疏漏与不足之处，敬请读者朋友指正。

